

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 23:29:42 ; Search time 1686 Seconds
(without alignments)
10093.953 Million cell updates/sec

Title: US-10-060-069-705
Perfect score: 416
Sequence: 1 gacaaatctgtaacctgacg.....gcatgcttccagcgctgct 416

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	356	85.6	142436	10	AL831763	AL831763 Mouse DNA
2	282.4	67.9	245136	2	AC098286	AC098286 Rattus no
3	221.4	53.2	864	6	AX358339	AX358339 Sequence
4	221.4	53.2	864	6	AX380557	AX380557 Sequence
5	221.4	53.2	1677	6	AX463507	AX463507 Sequence
C 6	221.4	53.2	115515	9	H5J1056H1	AL118505 Homo DNA
7	219.8	52.8	2731	9	BC027720	BC027720 Homo sapi
8	216.6	52.1	2223	6	AX358342	AX358342 Sequence
9	164.2	39.5	861	6	AX358341	AX358341 Sequence
10	159.4	38.3	2220	6	AX358344	AX358344 Sequence
11	70	16.8	5512	3	AF444780	AF444780 Anopheles
C 12	66.6	16.0	193660	2	AC102262	AC102262 Mus muscu
13	64.6	15.5	2125	6	AX409766	AX409766 Sequence
14	64.6	15.5	2125	9	HUMIGPACID	M86826 Human IGF b
15	64.6	15.5	2168	9	BC025681	BC025681 Homo sapi
16	64.6	15.5	4574	9	AF192554	AF192554 Homo sapi
C 17	64.6	15.5	21419	9	HS44786	AL031724 Human DNA
C 18	64.6	15.5	69437	9	AC012180	AC012180 Homo sapi
19	64.4	15.5	1818	9	S83462	S83462 Papio hamad
20	63.8	15.3	1416	6	AX392985	AX392985 Sequence
21	63.8	15.3	1507	6	AX392993	AX392993 Sequence
22	63.8	15.3	1880	9	AK027100	AK027100 Homo sapi
23	63.8	15.3	3582	6	AX704768	AX704768 Sequence
C 24	63.8	15.3	214530	9	AC020916	AC020916 Homo sapi
25	63	15.1	3159	6	E38228	E38228 MASL1 gene
26	63	15.1	3159	9	AB016816	AB016816 Homo sapi
27	63	15.1	6125	6	AX188327	AX188327 Sequence
28	63	15.1	6242	6	E38229	E38229 MASL1 gene
C 29	63	15.1	182622	9	AC090567	AC090567 Homo sapi
30	63	15.1	185510	2	AC011018	AC011018 Homo sapi
31	62.8	15.1	3634	10	BC048381	BC048381 Mus muscu
32	62.8	15.1	3635	10	BC035244	BC035244 Mus muscu
33	62.8	15.1	5318	10	AY114303	AY114303 Mus muscu
34	62.8	15.1	209799	10	AL671117	AL671117 Mouse DNA
35	62.4	15.0	141671	10	AC129082	AC129082 Mus muscu
C 36	62.4	15.0	146143	10	AC121896	AC121896 Mus muscu
C 37	61.2	14.7	190952	2	AC119020	AC119020 Rattus no
38	61.2	14.7	259561	2	AC116095	AC116095 Rattus no
39	61	14.7	442	4	AF218917	AF218917 Sus scro
40	60.6	14.6	1562	6	AX430385	AX430385 Sequence
41	60.6	14.6	1811	6	AX465479	AX465479 Sequence
42	60.6	14.6	1860	6	AX497238	AX497238 Sequence
C 43	60.6	14.6	2026	9	AK091795	AK091795 Homo sapi
44	60.6	14.6	2222	6	AX392971	AX392971 Sequence
45	60.6	14.6	2300	6	AX463719	AX463719 Sequence

ALIGNMENTS

RESULT 1
AL831763/c
LOCUS AL831763 142436 bp DNA linear ROD 24-OCT-2002
DEFINITION Mouse DNA sequence from clone RP23-308A23 on chromosome 2, complete
sequence.
ACCESSION AL831763
VERSION AL831763.5 GI:24395128
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 142436)
AUTHORS Leongamornlert, D.
TITLE Direct Submission

|||||
670 GCCTGAGCAGCCTCCGCGCTCGCTCGCGGAATCGCTCGCGGCGCTCAGCC 729
QY 388 CGGACCTTCATCTCCCGCCTGC 415
Db 730 CGGGCCTTCGCTCTCCCGCGCTGC 757

RESULT 6
HSJ1056H1/c
LOCUS
DEFINITION
Human DNA sequence from clone RP5-1056H1 on chromosome 20. Contains the gene for a putative novel leucine rich protein, the gene for a novel protein similar to mitogen inducible protein MIG-2, ESTs, STSS, GSSs and two CpG islands, complete sequence.
AL118505
ALL18505.17 GI:10639090
KEYWORDS
HTG; CpG island; MIG-2.
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 115515)
Direct Submission
Submitted (02-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Oct 4, 2000 this sequence version replaced gi:9795214.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
IMPORTANT: This sequence is not the entire insert of clone RP5-1056H1 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true right end of clone RP5-1056H1 is at 115515 in this sequence. The true left end of clone RP11-318P23 is at 84826 in this sequence. The true right end of clone RP5-967N21 is at 100 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-1056H1 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES
source
1. 115515
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP5-1056H1"
/clone.lib="RPCI-5"
39..371
repeat_region
/note="AluJo repeat: matches 1..289 of consensus"
343..431
repeat_region
/note="MLT1H repeat: matches 1..87 of consensus"

repeat_region
1349..1437
/note="Charlie4a repeat: matches 232..327 of consensus"
1438..1491
/note="9 copies 6 mer cacaca 94% conserved"
1516..1813
/note="AlusC repeat: matches 1..292 of consensus"
2093..2458
/note="L2 repeat: matches 1332..1724 of consensus"
2461..2675
/note="MIR repeat: matches 8..230 of consensus"
2737..3021
/note="L2 repeat: matches 1420..1746 of consensus"
3116..3400
/note="L2 repeat: matches 2357..2627 of consensus"
3405..3464
/note="10 copies 6 mer tcataca 71% conserved"
3506..3813
/note="Alusq repeat: matches 1..306 of consensus"
3908..3979
/note="12 copies 6 mer atcatc 76% conserved"
3999..4115
/note="MIR repeat: matches 121..235 of consensus"
complement(5198..5883)
/note="match: GSS: Em:AQ455346"
5212..5309
/note="MIR repeat: matches 23..123 of consensus"
5811..5958
/note="MIR repeat: matches 3..173 of consensus"
6041..6882
/note="CpG island"
evidence=not_experimental
complement(6264..6714)
/gene="dJ1056H1.1"
complement(<6264..>6714)
/gene="dJ1056H1.1"
/codon_start=1
evidence=not_experimental
/product="dJ1056H1.1 (putative leucine rich protein)"
/protein_id="CAC14960.1"
/db_xref="GI:13811937"
translation="GGPWSSGSNATSPCEGLPAADATALTANRNLERLPGLPRT
LRSDASHNLLRALSTSELGLEQLVLTLRHNRAALRWGPGFAGLHTLDLSYNQL
AALPCTCPALSSLRALALAGNLRALQPRAFACFPALQLNLCTAL"
8708..9066
/note="THE1A repeat: matches 1..354 of consensus"
9191..9222
/note="16 copies 2 mer tt 87% conserved"
9827..9955
/note="FLAM_C repeat: matches 1..140 of consensus"
10578..10688
/note="MIR repeat: matches 106..245 of consensus"
10745..11048
/note="Alusx repeat: matches 1..299 of consensus"
11557..11890
/note="L1R16C repeat: matches 63..382 of consensus"
12085..12180
/note="L2 repeat: matches 2636..2747 of consensus"
12619..12668
/note="MIR repeat: matches 81..130 of consensus"
13767..13899
/note="L2 repeat: matches 2575..2701 of consensus"
13944..14050
/note="L2 repeat: matches 2589..2699 of consensus"
14051..14349
/note="Alusq repeat: matches 1..300 of consensus"
14629..14805
/note="L1ME3 repeat: matches 5344..5518 of consensus"
14813..14877
/note="FLAM repeat: matches 60..123 of consensus"
14884..15303
/note="L1ME3 repeat: matches 5519..5987 of consensus"
15554..15939
/note="MLT1I repeat: matches 1..410 of consensus"

repeat_region	15972..16364	/note="L2 repeat: matches 2366. .2750 of consensus"
repeat_region	16366..16572	/note="L2 repeat: matches 1929. .2151 of consensus"
repeat_region	16573..16738	/note="L2 repeat: matches 1. .153 of consensus"
repeat_region	16739..17053	/note="L2 repeat: matches 1. .313 of consensus"
repeat_region	17054..17128	/note="L2 repeat: matches 153. .230 of consensus"
repeat_region	17132..17360	/note="L2 repeat: matches 1. .251 of consensus"
repeat_region	17376..17860	/note="L2 repeat: matches 1175. .1746 of consensus"
repeat_region	17897..17951	/note="L2 repeat: matches 477. .533 of consensus"
repeat_region	17952..18155	/note="L2 repeat: matches 404. .605 of consensus"
repeat_region	18156..18457	/note="L2 repeat: matches 1. .301 of consensus"
repeat_region	18458..18852	/note="L2 repeat: matches 1. .404 of consensus"
repeat_region	18853..19265	/note="L2 repeat: matches 77. .477 of consensus"
repeat_region	19289..19589	/note="L2 repeat: matches 1. .302 of consensus"
repeat_region	19654..19966	/note="L2 repeat: matches 744. .1123 of consensus"
repeat_region	20521..20544	/note="L2 repeat: matches 1. .252 of consensus"
repeat_region	21004..21094	/note="L2 repeat: matches 432. .547 of consensus"
repeat_region	21499..21799	/note="L2 repeat: matches 167. .466 of consensus"
repeat_region	22503..22611	/note="L2 repeat: matches 1. .294 of consensus"
repeat_region	22615..22785	/note="L2 repeat: matches 6. .167 of consensus"
repeat_region	23247..23498	/note="L2 repeat: matches 97. .175 of consensus"
repeat_region	23677..23803	/note="L2 repeat: matches 9. .38 of consensus"
repeat_region	24252..24532	/note="L2 repeat: matches 211. .278 of consensus"
repeat_region	24533..24827	/note="L2 repeat: matches 1. .127 of consensus"
repeat_region	24828..24994	/note="L2 repeat: matches 1. .310 of consensus"
repeat_region	25218..25296	/note="L2 repeat: matches 1. .299 of consensus"
repeat_region	25319..25348	/note="L2 repeat: matches 7. .313 of consensus"
repeat_region	25907..25974	/note="L2 repeat: matches 53.2%; Score 221.4; DB 9; Length 115515;
repeat_region	26469..26593	/note="L2 repeat: matches 74.0%; Pred. No. 7.4e-29;
repeat_region	26677..26987	/note="L2 repeat: matches 0; Mismatches 100; Indels 1; Gaps 1;
misc_feature	complement(27209..27716)	
repeat_region	27316..27605	
repeat_region	27774..28084	
repeat_region	/note="L2 repeat: matches 1. .299 of consensus"	
repeat_region	/note="L2 repeat: matches 7. .313 of consensus"	
Query Match	53.2%; Score 221.4; DB 9; Length 115515;	
Best Local Similarity	74.0%; Pred. No. 7.4e-29;	
Matches 287; Conservative	0; Mismatches 100; Indels 1; Gaps 1;	
QY	28 CTGACTCTGTATGAGNANGACTTCCCGGNGGAGTCAACCTCTGACTTGTGNNAC 87	
DB	6680 CGACTCCCTCGGAGGGTGCCTCCGCGGATGCGAGCGCTTACCTTGCAGC 6622	

QY	88 CCGNGGGGAGAGGCTGCCTAGCTGCCTGCGCGCAGCTGCGAGCCTCGATGCGACG 147	
DB	6621 CGCAACCTGGAGCGCTGCCTGCGCGCTTACCGCGCACACTGCGAGCTCGACGCGACG 6562	
QY	148 CACAACCTGCTGCGCGCTTACCGAGCTGTGCTGCGCGCTGCGCGCTGCGCGCTA 207	
DB	6561 CACAACCTGCTGCGCGCTTACCGAGCTGTGCTGCGCGCTGCGCGCTGCGCGCTG 6502	
QY	208 CTCACGCTGACCAACCGCATCTCCGCTTACACTGGGGCGCTGACACGCTGCGCGAG 267	
DB	6501 CTGACCTGCGCCACACCGCATCGCGCTGCGCTGCGCGCTGCGCGCTGCGCGG 6442	
QY	268 CTGCTGAGCTCGACCTCAGTCACACCTGCTGACGAGCTGCGCGCTGCGCGGCGCC 327	
DB	6441 CTGACACCTCGACCTCAGTCACACCTGCTGACGAGCTGCGCGCTGCGCGCTG 6382	
QY	328 TCGGGGAGAGCTGCGCTGCGCTGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGC 387	
DB	6381 GCGCTGAGCAGCTGCGCGCTGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCT 6322	
QY	388 CGGACCTTGTGCTGCTTCCGAGCGCTGC 415	
DB	6321 CGGCGCTTGTGCTGCTTCCGCGCTGC 6294	
RESULT 7		
BC027720		
LOCUS	Homo sapiens chromosome 20 open reading frame 75, mRNA (CDNA clone	
DEFINITION	MGC:20770 IMAGE:4588586), complete cds.	
ACCESSION	BC027720	
VERSION	BC027720.1 GI:20381180	
KEYWORDS	MGC.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 2731)	
REFERENCES	Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheet, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.	
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
MEDLINE	22388257	
PUBMED	12477932	
REFERENCE	2 (bases 1 to 2731)	
AUTHORS	Straussberg, R.	
TITLE	Direct Submission	
JOURNAL	Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov	
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Rubin Laboratory	

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mc@nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgon, C., Vost, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 30 Row: 1 Column: 14.

FEATURES

source

1. .2731
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:20770 IMAGE:458856"
 /tissue="Kidney, renal cell adenocarcinoma"
 /clone_lib="NIH_MGC_14"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1. .2731
 /gene="C20orf75"
 /note="synonyms: MGC25027, dJ1056H1.1"
 /db_xref="LocusID:164312"
 216. .2438
 /codon_start=1
 /product="chromosome 20 open reading frame 75"
 /protein_id="AAH27720.1"
 /db_xref="GI:20381181"
 /db_xref="LocusID:164312"
 /translation="MRQTLPLLLTLVLRSPWADPQEKVPLFRVTVQGPWGSSGNSAT
 DSPCEGLPADATALTANRNLRLPGCLPTLRLSLDASHNLLRALSTSELGHLEQLQ
 VLTLRHNRIFAALRWGPGAGLHLDLSYNOLAAALPCTGPALSLRALAGNPLRA
 LOAFACFPALQLNLSCALGRAGAGGIAEAFAGEDGAPVLTLEVLDSLGFLEFLER
 TPIQLVLFQNCLSLSPFPPMTLDSQVLSINLGNPLTSCDLSWLLTDARKTVLSRA
 ADTWCAPAGSGGPFSAISLQPLGVCOSDSTTLGASHPPCFNRSTYAGTVAPRAA
 AAPATRAGDOQSVSKAPNVGSRTIAAWPHSDAREGTAPSTNSVAGHSNSVFPRAA
 STTTRQHRGEAPLVLEPDISAASPLAKLLGFPFTSWDRSISSPQPGQRTHTPQ
 APNPSLSEGEIPVLLDDYSEEEGRKEEVTGPHDVPDYPHCKHLOTCAEALQRRW
 RCPCGLSGEDTIPDPRLQGVETTTTSALVHWCAPNSVVGQIIRYSAGFWAGNOS
 VVGVIYATARQHPLYGLSPGTYTYRVCVLAANRAGLSQPRSSGWRSPCAAFTKPSFAL
 LLSGLCAASGULLASTVLSACLRCRRGQTLGLQRCDTHLVAYKNPAFDYPLGLQTVS

gene

CDS

1. .2223
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 1. .2223
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD23325.1"
 /db_xref="GI:18674979"
 /translation="MRQTLPLLLTLVLRSPWADPQEKVPLFRVTVQGPWGSSGNSAT
 DSPCEGLPADATALTANRNLRLPGCLPTLRLSLDASHNLLRALSTSELGHLEQLQ
 VLTLRHNRIFAALRWGPGAGLHLDLSYNOLAAALPCTGPALSLRALAGNPLRA
 LOAFACFPALQLNLSCALGRAGAGGIAEAFAGEDGAPVLTLEVLDSLGFLEFLER
 TPIQLVLFQNCLSLSPFPPMTLDSQVLSINLGNPLTSCDLSWLLTDARKTVLSRA
 ADTWCAPAGSGGPFSAISLQPLGVCOSDSTTLGASHPPCFNRSTYAGTVAPRAA
 AAPATRAGDOQSVSKAPNVGSRTIAAWPHSDAREGTAPSTNSVAGHSNSVFPRAA
 STTTRQHRGEAPLVLEPDISAASPLAKLLGFPFTSWDRSISSPQPGQRTHTPQ
 APNPSLSEGEIPVLLDDYSEEEGRKEEVTGPHDVPDYPHCKHLOTCAEALQRRW
 RCPCGLSGEDTIPDPRLQGVETTTTSALVHWCAPNSVVGQIIRYSAGFWAGNOS
 VVGVIYATARQHPLYGLSPGTYTYRVCVLAANRAGLSQPRSSGWRSPCAAFTKPSFAL
 LLSGLCAASGULLASTVLSACLRCRRGQTLGLQRCDTHLVAYKNPAFDYPLGLQTVS

BASE COUNT

ORIGIN

Query Match 52.8%; Score 219.8; DB 9; Length 2731;
 Best Local Similarity 73.7%; Pred. No. 3.5e-28;
 Matches 286; Conservative 0; Mismatches 101; Indels 1; Gaps 1;
 QY 28 CTGACTCTGTATGAGNANGACTNCCCGGNGGGAGTACACACTCTGAACCTGGNNAAC 87
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 DB 346 CCGACTCCCTGCGAGGGGCTGCCCGC-CGCGGATGCGAGCGCTTGACCTGCGGAGC 404
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 QY 88 CGGNGGNGGAGAGCGCTGCCTAGCTGCTCGCGGCGAGCTGCGAGCGCTCGATGGCAGC 147
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 405 CGCAACTGGAGCGCTGCCCGGCTGCTACCGGCGACACTGCGAGCTGCGAGCGCGCAGC 464
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 148 CACAACCTGTCGCGCGCCCTTAGCGAGCGCTGTGCTCGCGCGCTCGCCGAGCTGCGCGTA 207
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 465 CACAACCTGTCGCGCGCCCTTAGCGACTTCCGAGCTCGGCCACCTGCGAGCGCTGCGAGTG 524
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 208 CTCACGCTGCACCAACCGCATCTCCGTGCTACACTGGGCGCGTGCACACGCTGCGCGAG 267
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 525 CTGACCTCTGCCACAAACCGCATGCCGCTGCGCTGGGCGCGGCTGGCGCGGGG 584
 QY 268 CTGCGTAGCTCGACCTCAGTCAACACTGCTGACGAGCTGCGCGGCTGCGCGGGGCC 327
 ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 585 CTGCACACCTTGGACCTCAGCTACACAGCTGGCCGCTGCTGCTGCGCGGCGCTGACGGGCC 644
 QY 328 TCGGGAGAGCAGCTGCGCTGCTGCGGCGGGAACCCACTGCGAGCGCTGCTGCCA 387
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 645 CGCTGTAGCAGCTCCGCGCCCTGCGCTGCGCGGGAATCCGCTGCGGGCGCTGACAGGCC 704
 QY 388 CGGACCTTTCATGCTTCCAGCGCTGC 415
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 705 CGGCGCTTCGCTGCTTCCCGCGCTGC 732

RESULT 8

AX358342

LOCUS

AX358342

DEFINITION

Sequence 13 from Patent WO0202604.

AX358342

ACCESSION

AX358342.1

VERSION

GI:18674978

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1

Thayer, E.C., Sheppard, P.O. and Presnell, S.R.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Leucine-rich repeat proteins, zlr7, zlr8 and zlr9

JOURNAL

Patent: WO 0202604-A 13 10-JAN-2002;

ZymoGenetics, Inc. (US)

FEATURES

Location/Qualifiers

1. .2223

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

1. .2223

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD23325.1"

/db_xref="GI:18674979"

/translation="MRQTLPLLLTLVLRSPWADPQEKVPLFRVTVQGPWGSSGNSAT
 DSPCEGLPADATALTANRNLRLPGCLPTLRLSLDASHNLLRALSTSELGHLEQLQ
 VLTLRHNRIFAALRWGPGAGLHLDLSYNOLAAALPCTGPALSLRALAGNPLRA
 LOAFACFPALQLNLSCALGRAGAGGIAEAFAGEDGAPVLTLEVLDSLGFLEFLER
 TPIQLVLFQNCLSLSPFPPMTLDSQVLSINLGNPLTSCDLSWLLTDARKTVLSRA
 ADTWCAPAGSGGPFSAISLQPLGVCOSDSTTLGASHPPCFNRSTYAGTVAPRAA
 AAPATRAGDOQSVSKAPNVGSRTIAAWPHSDAREGTAPSTNSVAGHSNSVFPRAA
 STTTRQHRGEAPLVLEPDISAASPLAKLLGFPFTSWDRSISSPQPGQRTHTPQ
 APNPSLSEGEIPVLLDDYSEEEGRKEEVTGPHDVPDYPHCKHLOTCAEALQRRW
 RCPCGLSGEDTIPDPRLQGVETTTTSALVHWCAPNSVVGQIIRYSAGFWAGNOS
 VVGVIYATARQHPLYGLSPGTYTYRVCVLAANRAGLSQPRSSGWRSPCAAFTKPSFAL
 LLSGLCAASGULLASTVLSACLRCRRGQTLGLQRCDTHLVAYKNPAFDYPLGLQTVS

BASE COUNT 374 a 824 c 663 g 362 t
 ORIGIN
 Query Match 52.1%; Score 216.6; DB 6; Length 2223;
 Best Local Similarity 73.2%; Pred. No. 1.3e-27;
 Matches 284; Conservative 0; Mismatches 103; Indels 1; Gaps 1;
 QY 28 CTGACTCTGTATGAGNANGACTNCCCGGNGGGAGTACACACTCTGAACCTGGNNAAC 87
 ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 131 CCGACTCCCTGCGAGGGGCTGCCCGC-CGCGGATGCGAGCGCTTGACCTGCGGAGC 189
 ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 88 CGGNGGNGGAGAGCGCTGCCTAGCTGCTCGCGGCGAGCTGCGAGCGCTCGATGGCAGC 147
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 190 CGCAACTGGAGCGCTGCCCGGCTGCTACCGGCGACACTGCGAGCGCTGCGAGCGCAGC 249
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 148 CACAACCTGTCGCGCGCCCTTAGCGAGCGCTGTGCTCGCGCGCTCGCCGAGCTGCGCGTA 207
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 250 CACAACCTGTCGCGCGCCCTTAGCGACTTCCGAGCTGCGGCCACCTGCGAGCGCTGCGAGTG 309
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||


```

Db      803 GGACCGCAACCTCATCGCTGCGGTGGCCCGGCGCCCTCTCTGGGCGCTGAAGGCGCTGCG 862
Qy      273 TGAGCTGACCTCAGTACACACCTGCTGACGAGCTGCGCGCCCTCGCGGGCGCCCTCGGG 332
Db      863 ATGGCTGACCTGTCCACACCGCGGTGGCTGCTCTGGAGACAGCTTCCCGGCTCT 922
Qy      333 GAGCAGCGCTGCGCTGCGGTGGCGGGGACCACTGCGAGCGCTGCTGCCACGCGAC 392
Db      923 GCTGGCGCTGCTGTGCTGCGGCTGTCCACACGCGCATGCCAGCTGCGGCGCCGCGAC 982
Qy      393 CTT 395
Db      983 CTT 985

RESULT 14
LOCUS   HUMIGFACID
DEFINITION Human IGF binding protein complex acid-labile subunit a mRNA,
complete cds.
ACCESSION M86826
VERSION 1
KEYWORDS acid-labile subunit; insulin-like growth factor binding protein.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2125)
AUTHORS Leong,S.K., Baxter,R.C., Camerato,T., Dai,J. and Wood,W.I.
TITLE Structure and functional expression of the acid-labile subunit of
the insulin-like growth factor-binding protein complex
JOURNAL Mol. Endocrinol. 6 (6), 870-876 (1992)
MEDLINE 92357025
PUBMED 1379671
COMMENT Original source text: Homo sapiens adult liver cDNA to mRNA.
FEATURES
source
location/Qualifiers
1..2125
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="liver"
/dev_stages="adult"
57..1874
/note="acid-labile subunits a"
/codon_start=1
/product="insulin-like growth factor binding protein
complex"
/protein_id="AAA36047.1"
/db_xref="GI:184808"
/translation="MALRKGGLALALLLSWVALGPSRLEGADPTGAEAGPACPA
CVCYDDADELSVFCSSRNLTPLDPGVGTQALWLDGNNLSSVPPAFAFNLSSLGF
LNLOGGOLGSLPEOALLGLENLCHLHLPNOLRLSLGTFHTPALASLGLSNNELSR
LEDGLFELGSLDNLNGNSLAVLPDPAFRGLGSLRELVLAGNRLAYLQALFSLGLA
ELREDLSRNALRAIKANVFOLPRLOKLYLDNRLLIAVAPGAFGLKALRDLDSHN
RVAGLEDTFFGLGLRVLRSLHNAIASLRLPTKDLHLEELQIGHNRIQLAERSF
EGLCQLEVLTDHQLQEVKAGAFGLTNVAMNLSNCLRNLPQVFRGLGKLHSLH
LEGSCGLRPHRTFTGLSRLFLKONGLVIEQSLWGLAELELDLTNSQTLHPL
HRLFGGLKLYLLSRNLAELPADALGPLORAFWLDVSNRLEALPNSLLAPLGL
EYLSLRNLSLFTFPQPPGLERLWEGNPWDCGCPKLKALRFALQNSAVPRFQAIC
EDDCQPPATYTNITCASPPVEVGLDLRLDUSEAHFAPC"
138..1871
mat_peptide
/product="insulin-like growth factor binding protein
complex"
/note="acid-labile subunits a"
/evidence="experimental"
BASE COUNT 389 a 752 c 657 g 327 t
ORIGIN
Query Match 15.5%; Score 64.6; DB 9; Length 2125;
Best Local Similarity 53.1%; Pred No. 0.11;
Matches 161; Conservative 0; Mismatches 139; Indels 3; Gaps 1;
Qy 96 GGAGAGCGCTGCTAGTGCCTGCCGCGCACGCTGCGCACGCTCGATGCGCACCAACCT 155

```

```

Db      683 GCAGCCGCGCTCTTACGCGCCTTGGCGAGCTCGGAGCTGGACCTGAGCAGGACGC 742
Qy      156 GCTGGCGCCCTTAGCGAGCCTGTCTCGCGCGCTGCGCGAGCTGCGCGTACTCACGCT 215
Db      743 GCTCGGGCGCATCAAGGCAACGTTCTGTCAGCTGCCCGGCTCCAGAACTACTACT 802
Qy      216 GCACCAACGCGATCTCCGTACACTGGGCGCGTGACACG---CTGGCGGAGCTGCG 272
Db      803 GGACCGCAACCTCATCGCTGCGGTGGCCCGGCGCCCTCTCTGGGCGCTGAAGGCGCTGCG 862
Qy      273 TGAGCTGACCTCAGTACACACCTGCTGACGAGCTGCGCGCCCTCGCGGGCGCCCTCGGG 332
Db      863 ATGGCTGACCTGTCCACACCGCGGTGGCTGCTCTGGAGACAGCTTCCCGGCTCT 922
Qy      333 GAGCAGCGCTGCGCTGCGGTGGCGGGGACCACTGCGAGCGCTGCTGCCACGCGAC 392
Db      923 GCTGGCGCTGCTGTGCTGCGGCTGTCCACACGCGCATGCCAGCTGCGGCGCCGCGAC 982
Qy      393 CTT 395
Db      983 CTT 985

BC025681 2168 bp mRNA linear PRI 11-MAR-2002
Homo sapiens, insulin-like growth factor binding protein, acid
labile subunit, clone MGC:34474 IMAGE:5205012, mRNA, complete cds.
ACCESSION BC025681
VERSION BC025681.1 GI:19344009
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2168)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo..
TITLE Direct Submission
JOURNAL Strausberg,R.
Submitted (06-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nigri.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantropop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 49 Row: g Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4826771.
Location/Qualifiers
1..2168
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:3483"
/db_xref="taxon:9606"
FEATURES
source

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 23:34:31 ; Search time 250 Seconds
(without alignments)
4491.865 Million cell updates/sec

Title: US-10-060-069-705
Perfect score: 416
Sequence: 1 gacaaatctctacgtgacg.....gcattgtcccgagcgctgct 416

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	223.6	53.8	3665	22 AAD05513	Human secreted pro
2	222.8	53.6	1551	22 AAD05562	Human secreted pro
3	222.8	53.6	1590	24 ABQ54274	Human ovarian anti
4	221.4	53.2	864	24 ABK33044	DNA encoding novel
5	221.4	53.2	864	24 AAD28126	Human leucine-rich
6	221.4	53.2	1054	24 AAD38697	Human LP222/LP222a
7	221.4	53.2	1677	24 ABK99934	DNA encoding human
8	216.6	52.1	2223	24 AAD28128	Human leucine-rich

9	164.2	39.5	861	24 AAD28127	Human leucine-rich
10	159.4	38.3	2220	24 AAD28129	Human leucine-rich
11	146.6	35.2	697	24 ABQ42438	Oligonucleotide fo
12	146.6	35.2	697	24 ABQ42439	Oligonucleotide fo
13	83.2	20.0	697	24 ABQ42440	Oligonucleotide fo
14	83.2	20.0	697	24 ABQ42441	Oligonucleotide fo
15	64.6	15.5	2125	24 ABK95915	Gene #2413 used to
16	63.8	15.3	1416	24 ABK62088	Human cDNA encodin
17	63.8	15.3	1452	22 AAI67182	Nucleotide sequenc
18	63.8	15.3	1507	24 ABK62092	Human cDNA encodin
19	63.8	15.3	1818	22 AAI67183	Nucleotide sequenc
20	63	15.1	3159	21 AAA99494	Human DNA encoding
21	63	15.1	3400	21 AAC76598	Human ORFX ORF2153
22	63	15.1	6125	22 AAH72745	Human cervical can
23	63	15.1	6242	21 AAA99495	Human cDNA encodin
24	60.6	14.6	1562	24 ABQ72615	Human MDDT encodin
25	60.6	14.6	1779	22 AAD16344	Human sbgtango79a
26	60.6	14.6	1779	24 ABQ86161	DNA encoding human
27	60.6	14.6	1811	24 ABK48387	DNA encoding novel
28	60.6	14.6	1860	24 ABK92037	Human cDNA encodin
29	60.6	14.6	2222	24 ABK62081	Human LP polypepti
30	60.6	14.6	2300	24 AAD41275	Human CGDD cDNA 68
31	60.6	14.6	2338	25 ABQ77432	Human extracellula
32	60.4	14.5	2390	22 AAD08046	Human ovarian anti
33	60.4	14.5	3125	24 ABQ54988	Protein PRO332 cDN
34	60.4	14.5	3296	20 AAX52267	Human PRO332 prote
35	60.4	14.5	3296	21 AAC58590	Human cDNA sequenc
36	60.4	14.5	3296	22 AAS21428	Human cDNA sequenc
37	60.4	14.5	3296	22 AAF72425	Human PRO332 cDNA.
38	60.4	14.5	3296	25 ACA55065	Novel human secret
39	60.4	14.5	3296	25 ACA58549	cDNA encoding huma
40	60.4	14.5	3296	25 ACA60256	Human cDNA for sec
41	60.4	14.5	3296	25 ACA03787	cDNA encoding huma
42	60.4	14.5	3296	25 ACA04208	Human cDNA encodin
43	60.4	14.5	3296	25 ACA05594	cDNA encoding huma
44	60.4	14.5	3296	25 ABX89325	DNA encoding novel
45	60.4	14.5	3296	25 ABX96273	Human secreted/tra

ALIGNMENTS

RESULT 1

AAD05513

ID AAD05513 standard; cDNA; 3665 BP.

XX AAD05513;

AC AAD05513;

XX 18-JUL-2001 (first entry)

DE Human secreted protein-encoding gene 22 cDNA clone HOFNM53, SEQ ID NO:32.

XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma; foetal abnormality; developmental abnormality; haematopoietic disorder; immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis; Parkinson's disease; cognitive disorder; schizophrenia; skin disorder; psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder; inflammation; neurological disorder; Alzheimer's disease; food additive; anglogenetic disorder; kidney disorder; gastrointestinal disorder; allergy; pregnancy-related disorder; endocrine disorder; infection; wound healing; cell culture; chemotaxis; vulnery; binding partner identification; gene therapy; ss.

OS Homo sapiens.

XX	Key	Location/Qualifiers
FT	CDS	208..2442
FT		/*tag= a
FT		/product= "Human secreted protein precursor"
FT	sig_peptide	208..270
FT	mat_peptide	271..2439
FT		/*tag= b
FT		/*tag= c

FT XX /product- "Mature human secreted protein"

PN WO200134767-A2.

XX 17-MAY-2001.

XX 01-NOV-2000; 2000WO-US30036.

XX 05-NOV-1999; 99US-0163576.

PR 27-JUL-2000; 2000US-0221366.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Soppet DR, Komatsoulis G, Shi Y, Olsen HS, Ruben SM;

XX WPI: 2001-316492/33.

DR P-PSDB; AAE01693.

XX Isolated nucleic acid molecule encoding a human secreted protein is

PT used in preventing, treating or ameliorating a medical condition -

XX

PS Claim 1; Page 426-427; 540pp; English.

XX AAD05492-AA05564 represent cDNAs corresponding to 22 human secreted

CC protein genes, and AAE01743 represent the proteins they encode.

CC AAE01744-AAE01763 represent human secreted protein fragments or variants.

CC The secreted proteins and their genes are useful for preventing, treating

CC or ameliorating medical conditions, e.g., by protein or gene therapy.

CC Pathological conditions can be diagnosed by determining the amount of the

CC new protein in a sample or by determining the presence of mutations in

CC the new genes. Specific uses are described for each of the 22 genes,

CC based on the tissues in which they are most highly expressed, and include

CC developing products for the diagnosis or treatment of proliferative

CC disorders, cancer, tumours, foetal and developmental abnormalities,

CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune

CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,

CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),

CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,

CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,

CC angiogenic disorders, kidney disorders, gastrointestinal disorders,

CC pregnancy-related disorders, endocrine disorders, and infections. The

CC proteins can also be used to aid wound healing and epithelial cell

CC proliferation, to prevent skin aging due to sunburn, to maintain organs

CC before transplantation, for supporting cell culture of primary tissues,

CC to regenerate tissues, to identify their cognate ligands or binding

CC partners, and in chemotaxis, and can be used as a food additive or

CC preservative to modify storage properties. Antibodies specific for a

CC protein of the invention can be used in alleviating symptoms associated

CC with the disorders mentioned above, and in diagnostic immunossays e.g.,

CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).

CC The present sequence represents a human secreted protein-encoding cDNA of

CC the invention.

XX

SQ Sequence 3665 BP; 782 A; 1164 C; 967 G; 752 T; 0 other;

Query Match 53.8%; Score 223.6; DB 22; Length 3665;

Best Local Similarity 75.3%; Pred. No. 4e-35;

Matches 274; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 52 CCGCGNGGGAGTCACCACTTGAACTTGGNNAACCGGNGGNGAGAGCGCTGCCTAGC 111

DB 370 CCTCTGCGGGAGCGTCGACCTTGACCTGGGCGAACCGGAGCGCTGGAGCGCTGCCCAAC 429

QY 112 TGCTTCCCGCGGACGCTGGCGGCTGCGATGGCAGGCACAACTGCTGGCGGCGCTTAGC 171

DB 430 TGGCTCCCGCGCGCGCTGGGTAGCTTAGACGCCAGGCACAACTGCTGGCGGCGCTTAGC 489

QY 172 GAGCCTGTGCTCGGCGCGCTGCGCGAGCTGCGGCTACTCAGCGTCACCAACACCGCATC 231

DB 490 GTCCAGAGCTCGGCGCGCTGCGCGGCTGCGAGCGCTGACGCTGGCCACCAACCGCATC 549

QY 232 TCGGTGCTACCTGGGCGCGTCGACAGCGTGGCGGAGCTGCGTACCTGACCTCAGTCAC 291

Db 550 GCCGAGCTGCGCTGGGGCCCCCGGAGGGCGCGCGCTGCACACGCTGGACCTCAGCTAC 609

QY 292 AACCTGCTGACCGAGCTGCGCGCGCGCTGCGGGGGCCCTCGGGGAGAGAGCTGCGCTCGCTG 351

Db 610 AACGAGCTGGCCAGCTGCGCGCCATGCGCGGGCGCGCGCTGCCCGGCTCCGCTCGCTG 669

QY 352 CGCTGGCGGGGAACCCACTGCGAGCGCTGCTGCCACGGACCTTTGCATGCTTCCACAGCG 411

Db 670 GCGCTGCGGGGAACCCGCTGCGAGCGCTGCGAGCGCGCTGCGCGGCGCTTGCCTCCCTCCGCGCG 729

QY 412 CTGC 415

Db 730 CTGC 733

RESULT 2

AAD05562 standard; cDNA; 1551 BP.

XX AAD05562;

XX 18-JUL-2001 (first entry)

XX Human secreted protein-encoding gene 22 cDNA clone HOFNM53, SEQ ID NO:81.

DE Human; secreted protein; proliferative disorder; cancer; tumour;

KW foetal abnormality; developmental abnormality; haematopoietic disorder;

KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;

KW inflammation; allergy; neurological disorder; Alzheimer's disease;

KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;

KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;

KW cardiovascular disorder; angiogenic disorder; kidney disorder;

KW gastrointestinal disorder; pregnancy-related disorder; gene therapy;

KW endocrine disorder; infection; wound healing; vulnary;

KW cell culture; chemotaxis; food additive;

KW binding partner identification; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 217..315

FT /*tag= a

FT /product= "Human secreted protein precursor"

FT sig_peptide 217..270

FT /*tag= b

FT mat_peptide 271..312

FT /*tag= c

FT /product= "Mature human secreted protein"

XX WO200134767-A2.

PN 17-MAY-2001.

XX

PD 01-NOV-2000; 2000WO-US30036.

PF

XX 05-NOV-1999; 99US-0163576.

PR 27-JUL-2000; 2000US-0221366.

XX

XX (HUMA-) HUMAN GENOME SCI INC.

XX Soppet DR, Komatsoulis G, Shi Y, Olsen HS, Ruben SM;

PI WPI: 2001-316492/33.

DR P-PSDB; AAE01741.

XX

XX Isolated nucleic acid molecule encoding a human secreted protein is

PT used in preventing, treating or ameliorating a medical condition -

XX

XX Claim 1; Page 460-461; 540pp; English.

XX AAD05492-AA05564 represent cDNAs corresponding to 22 human secreted

CC protein genes, and AAE01743 represent the proteins they encode.

CC AAE01744-AAE01763 represent human secreted protein fragments or variants.

Db 417 CCCTCTGCGGAGCGKCGACCTTGACCCCTGGGGAACCGCAGCCTGGAGGCGCTGCCCAAC 476

QY 112 TGCTGCGCGGCGAGCTGCGGAGCCTCGATGGGAGCCACAACTGCTGGCGCCCTAGC 171

Db 477 TGCTGCGCGGCGGCTGCTGCTAGCCCTAGAGCGGAGCCACAACTGCTGGCGCCCTGAGC 536

QY 172 GAGCCTGTGCTGCGGCGCCTGCGCGGAGCTGCGGCTTACTCAGCTGCGGAGCCACAACTGCTG 231

Db 537 GCTCCAGAGCTGCGGCGCCTGCGCGGAGCTGCGGAGCCTGAGCGCTGCGGAGCCACAACTGCTG 596

QY 232 TCGCTGCTACACTGGGCGGCTGACACGCTGCGGAGCCTGCGTGGAGCTGACCTCAGTCAC 291

Db 597 GCGGAGCTGCTGCGGCGGCGGAGGCGGCGGCGCTGACACGCTGAGACCTCAGCTAC 656

QY 292 AACCTGCTGACGAGCTGCGGCGCCTGCGGCGGCGGAGGCGGCGGAGCCTGCGGAGCCTGCGT 351

Db 657 AACGAGCTGCGGAGCTGCGGCGCCTGCGGCGGCGGAGGCGGCGGAGCCTGCGGAGCCTGCGT 716

QY 352 GCGCTGCGGCGGAGCCTGCGGAGCCTGCGGAGCCTGCGGAGCCTGCGGAGCCTGCGGAGCCTGCGG 411

Db 717 GCGCTGCGGCGGAGCCTGCGGAGCCTGCGGAGCCTGCGGAGCCTGCGGAGCCTGCGGAGCCTGCGG 776

QY 412 CTGC 415

Db 777 CTGC 780

RESULT 4

ABK33044

ID ABK33044 standard; DNA; 864 BP.

XX AC ABK33044;

XX DT 08-MAY-2002 (first entry)

XX DE DNA encoding novel secreted protein 2835840G8P.

XX KW Protein secretion; mammalian secreted polypeptide; MSP;

XX KW gene; ss.

XX OS Homo sapiens.

XX PN WO200202621-A2.

XX PD 10-JAN-2002.

XX PF 28-JUN-2001; 2001WO-US20638.

XX PR 30-JUN-2000; 2000US-215446P.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Sheppard PO, Presnell SR;

XX PI WPI; 2002-147999/19.

DR P-PSDB; AAU83129.

XX Novel isolated mammalian secreted polypeptide useful in therapeutic and

PT diagnostic methods, to direct secretion of other proteins of interest

PT from host cell, as educational tools, and as laboratory practicum kits

PS Claim 3; Page 175-177; 397pp; English.

XX The invention describes an isolated mammalian secreted polypeptide (MSP)

CC (1). (1) is useful to direct the secretion of other proteins of interest

CC from a host cell, to monitor secretion of proteins, to degenerate

CC sequences comprising all nucleotide sequences encoding a particular

CC polypeptide, to screen for cell metabolism effecting receptors, for

CC identifying new target receptors and drug design, for identifying, for

CC protein purification, for determining the weight of expressed MSP

CC polypeptides as a ratio to total protein expressed, for identifying

CC peptide cleavage sites, for coupling amino and carboxy terminal tags, for

CC amino acid sequence analysis, for monitoring biological activities of the

CC protein in vitro and in vivo, and to teach analytical skills and as

CC reagents for the study of cells, receptors, and other binding molecules.

CC The polynucleotide is useful for radiation hybrid mapping, and somatic

CC cell genetic technique developed for constructing high-resolution,

CC contiguous maps of mammalian chromosomes. Reagents disclosed in the

CC invention may be used to detect metabolic abnormalities characterised by

CC over or under production of the protein. This sequence encodes an

CC mammalian secreted polypeptide, described in the method of the invention.

XX SQ Sequence 864 BP; 136 A; 327 C; 265 G; 136 T; 0 other;

Query Match 53.2%; Score 221.4; DB 24; Length 864;

Best Local Similarity 74.0%; Pred. No. 1.1e-34;

Matches 287; Conservative 0; Mismatches 100; Indels 1; Gaps 1;

QY 28 CTGACTCTGTATGAGNANGACTNCCCGGNGGGAGTACCACTCTGAACCTTGGNNAAC 87

Db 131 CGGACTCGCCCTGCGAGGGGCTGCCGC-CGCGGATCGGACGCTTACCTTGGCGAAC 189

QY 88 CGGNGGGGAGAGCCTGCTAGCTGCTGCGGCGCAGCTGCGGAGCTGCGATGCGCAGC 147

Db 190 CGCAACTGGAGGCGCTGCCGCTGCTTACCGCGCACACTGCGGAGCTTCCGACGCCAGC 249

QY 148 CACAACCTGCTGCGGCGCCTTACGAGCCTGCTGCGGCGCTGCGGAGCTGCGCGTA 207

Db 250 CACAACCTGCTGCGGCGCCTGAGCACTTCCGAGCTGCGGCACTTGGAGCAGCTGCGAGTG 309

QY 208 CTCACGCTGCACCAACCGCATCTCCGTGCTACATGCGGCGCTGACAGCTGCGCGAG 267

Db 310 CTACGCTGCGGCGCACACCGCATCTGCGGCTGCGGCGGCGGCGGCGGCGGCGG 369

QY 268 CTGCGTGAGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327

Db 370 CTGACACCTGGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429

QY 328 TCGGGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387

Db 430 GCGCTGAGCAGCTGCGGCGCCTGCGGCGCTGCGGCGGAAATCCGCTGCGGCGCTGCGAGCC 489

QY 388 CGGACCTTTCATGCTTCCGAGGCTGC 415

Db 490 CGGCGCTTCCGCTTCCGCGCTGC 517

RESULT 5

AAD28126

ID AAD28126 standard; DNA; 864 BP.

XX AC AAD28126;

XX DT 22-APR-2002 (first entry)

XX DE Human leucine-rich repeat-9 (ZLRR9) DNA #1.

XX KW Human; leucine-rich repeat-9; ZLRR9; cytostatic; gene therapy; leukaemia;

XX KW endometrial adenocarcinoma; renal cell; colon; prostate; retinoblastoma;

XX KW cell growth disorder; glioblastoma; neuroblastoma; kidney tumour; lung;

XX KW mammary gland; germ cell; bladder; oesophagus; pancreas; animal feed;

XX KW rhabdomyosarcoma; genitourinary tract; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 1..864

FT /tag= a

FT /product= "ZLRR9 protein"

FT sig_peptide 1..45

FT /tag= b

FT mat_peptide 45..861

FT /tag= c

FT misc_feature 46..141

FT /*tag= d
FT /note= "Encodes N-terminal Cys-Rich Region"
FT misc_feature 145..216
FT /*tag= e
FT /note= "Leu-Rich Region 1"
FT misc_feature 217..288
FT /*tag= f
FT /note= "Leu-Rich Region 2"
FT misc_feature 289..360
FT /*tag= g
FT /note= "Leu-Rich Region 3"
FT misc_feature 361..432
FT /*tag= h
FT /note= "Leu-Rich Region 4"
FT misc_feature 436..507
FT /*tag= i
FT /note= "Leu-Rich Region 5"
FT misc_feature 508..579
FT /*tag= j
FT /note= "Leu-Rich Region 6"
FT misc_feature 580..651
FT /*tag= k
FT /note= "Leu-Rich Region 7"
FT misc_feature 700..837
FT /*tag= l
FT /note= "Encodes C-terminal Cys-Rich Region"
FT misc_feature 883..1059
FT /*tag= m
FT /note= "Encodes IG domain"
FT misc_feature 1210..1461
FT /*tag= n
FT /note= "Encodes Fibronectin III domain"

WO200202604-A2.

10-JAN-2002.

02-JUL-2001; 2001WO-US20999.

30-JUN-2000; 2000US-215446P.

(ZYMO) ZYMOGENETICS INC.

Thayer EC, Sheppard PO, Presnell SR;

WPI; 2002-154725/20.

P-PSDB; AAE17485.

New leucine-rich repeat proteins and polynucleotides, useful for diagnosing and treating disorders related to abnormal cell growth e.g. retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia, kidney and lung tumors

Claim 47; Page 70-72; 82pp; English.

The invention relates to an isolated polypeptide comprising leucine-rich repeat proteins Zlrr7, Zlrr8, Zlrr9 and spliced variants of Zlrr8, Zlrr9. Zlrr7, Zlrr8, and Zlrr9 proteins are useful in directing the secretion of proteins of interest from a host cell and to monitor the secretion of proteins in general from cells and tissues. The Zlrr DNA and proteins are useful in diagnosing and treating disorders related to abnormal cell growth, including retinoblastoma, renal cell adenocarcinoma, endometrial adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphocytic leukemia, kidney tumors, germ cell tumors, lung large cell carcinoma, mammary, colon adenocarcinoma, genitourinary tract transitional cell tumours, rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas and prostate adenocarcinoma. Zlrr protein is useful for identifying agonists and antagonists of the polypeptide, for drug design, to screen for cell metabolism effecting receptors, for analysis of cell phenotype, and as animal feed supplement and cell culture components. Zlrr DNA is also useful in gene therapy. The present sequence is human Zlrr9 DNA.

Sequence 864 BP; 136 A; 327 C; 265 G; 136 T; 0 other;

Query Match 53.2%; Score 221.4; DB 24; Length 864;
Best Local Similarity 74.0%; Pred. No. 1.e-34;
Matches 287; Conservative 0; Mismatches 100; Indels 1; Gaps 1;
QY 28 CTGACTCTGTATGAGNANGACTNCCGCGNGNGGAGTGACACACTCTGAACCTTGGNNAAC 87
DB 131 CCGACTCGCCCTGGGAGGGGCTGCCCGC-CGCGGATGCGAGGGCTTGACCTTGCGGAAAC 189
QY 88 CGGNGGNGGAGAGCGCTGCTAGCTGCTGCGCGGACGCTGCGGAGCTCGATGGCAGC 147
DB 190 CGCAACCTGGAGCGCTGCCCGCTGCTACCGCGCACACTGCGGAGCTTGACCCGCA 249
QY 148 CACAACCTGCTGCGCGCCCTTAGCGAGCTGTGCTGCGCGCGCTGCCGAGCTGGCGCTA 207
DB 250 CACAACCTGCTGCGCGCCCTTAGCGACTTCCGAGCTGCGCCACCTGGAGCAGCTGCGAGTG 309
QY 208 CTCAGCTGTCACACACACCGGATCTCGTGTACACTGGGGCGGTGACACGCTGGCCGAG 267
DB 310 CTGACCTTGGCGCCACACGCGATCGCGCTGCGCTGGGGCCCGGGTGGCGCGGGG 369
QY 268 CTGCGTGAGCTGCGACCTCAGTCACAACTGCTGACCGAGCTGCGCGCGCTGGCGGGGCC 327
DB 370 CTGCACACCTTGGAGCTTACAGCTACAACTGCGCGCTGCGCGCGCTGCGCGGGGCC 429
QY 328 TCGGGGAGCAGCTGCGCTGCTGCGCTGCGCGGGAACCCACTGCGAGCGCTGCTGCCA 387
DB 430 CGCTGAGCAGCTTCCGCGCTGCGCTGCGCGGGAATCCGCTGCGGGCGCTGCGAGCCC 489
QY 388 CGGACCTTTGCTATGCTTCCGAGCGCTGC 415
DB 490 CGGGCTTTCGCTGCTTCCCGCGCTGC 517

RESULT 6

AAD38697

ID AAD38697 standard; cDNA; 1054 BP.

AC AAD38697;

DT 23-SEP-2002 (first entry)

DE Human LP222/LP222a secreted protein encoding cDNA.

XX Human; secreted protein; atherosclerosis; Alzheimer's disease; LP222;
XX diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
XX rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
XX reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
XX gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;
XX chromosome 20; LP222a; gene; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT 119..1054

FT /*tag= a

FT /product= "Human LP222/LP222a secreted protein"

FT sig_peptide 119..172

FT /*tag= b

FT mat_peptide 173..1051

FT /*tag= c

FT /product= "Mature human LP222/LP222a secreted protein"

XX WO200226801-A2.

PN 04-APR-2002.

XX 14-SEP-2001; 2001WO-US26026.

XX 28-SEP-2000; 2000US-236088P.

XX (ELIL) LILLY & CO ELI.

XX

PF 02-JUL-2001; 2001WO-US20999.
XX
PR 30-JUN-2000; 2000US-215446P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Thayer EC, Sheppard PO, Presnell SR;
XX
XX WPI; 2002-154725/20.
XX
XX New leucine-rich repeat proteins and polynucleotides, useful for
PT diagnosing and treating disorders related to abnormal cell growth e.g.
PT retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia,
PT kidney and lung tumors
XX
XX Disclosure; Page 81-82; 82pp; English.
XX
XX The invention relates to an isolated polypeptide comprising leucine-rich
CC repeat proteins Zlrr7, Zlrr8, Zlrr9 and spliced variants of Zlrr8, Zlrr9.
CC Zlrr7, Zlrr8, and Zlrr9 proteins are useful in directing the secretion of
CC proteins of interest from a host cell and to monitor the secretion of
CC proteins in general from cells and tissues. The Zlrr DNA and proteins
CC are useful in diagnosing and treating disorders related to abnormal cell
CC growth, including retinoblastoma, renal cell adenocarcinoma, endometrial
CC adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphocytic leukemia,
CC kidney tumours, germ cell tumours, lung large cell carcinoma, mammary,
CC colon adenocarcinoma, genitourinary tract transitional cell tumours,
CC rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas,
CC and prostate adenocarcinoma. Zlrr protein is useful for identifying
CC agonists and antagonists of the polypeptide, for drug design, to screen
CC for cell metabolism effecting receptors, for analysis of cell phenotype,
CC and as animal feed supplement and cell culture components. Zlrr DNA is
CC also useful in gene therapy. The present sequence is human Zlrr9
CC degenerate DNA.
XX
XX
SQ Sequence 2220 BP; 268 A; 325 C; 373 G; 236 T; 1018 other;
Query Match 38.3%; Score 159.4; DB 24; Length 2220;
Best Local Similarity 39.8%; Pred. No. 1.4e-22;
Matches 148; Conservative 74; Mismatches 150; Indels 0; Gaps 0;
QY 44 ANGACTNCCCGGNGGAGTCAACACCTCTGACCTGGNNAACCGGNGGAGAGCC 103
DB 146 ARGGNYTNCNGCNGGAYGNACNCGNYTNACNYTNGCNAAYTMGNAAYTYNGARMGY 205
QY 104 TGCTAGCTGCTGCGCGGCGACGCTGCGAGCTGCTGATGCGACACACCTGCTCGCG 163
DB 206 TNCNGGNTGYTNCNMGACNYTNMGWSNYTNGAYGNCNNSNCAIAYTYNTNMGNG 265
QY 164 CCCTTAGCGAGCTGTGCTGCGCGGCTGCGGAGCTGCGTACTTACGCTGCACCA 223
DB 266 CNYTNWSNACNSNGARYTNGGNCAYTYNGARCARYTNACNYTNACNYTNMGNCAYA 325
QY 224 ACCGATCTCGTGTACTACTGGGCGGCGTACACGCTGCGGAGCTGCTGCTGCTGAC 283
DB 326 ATYGNATHGNCNTYNTMGTYTGGGNCNGGNGCNCNGNYTNGAYTNGAYTY 385
QY 284 TCAGTCACACCTGCTGACGAGCTGCGCGGCTGCGGCGGCGCTGCGGAGAGCTGCG 343
DB 386 TWSNTYAYARYCARYTNGCNGNYTNTCCNTGYACNGCNCNGCNYTNWSNYTNM 445
QY 344 GCTCGCTGCGTGGCGGGAACCCACTGCGAGCGCTGCTGCGCAGGCTTTGCGATGCT 403
DB 446 GNGCNYTNGCNYTNGCNGGNAAYCCNYTNMGNCNYTNACGNCNCNTTYGCTGYT 505
QY 404 TCCAGCGCTGC 415
DB 506 TYCCNGCNYTNC 517

RESULT 11
ABQ42438/c
ID ABQ42438 standard; DNA; 697 BP.

XX ABQ42438;
XX
XX 12-JUL-2002 (first entry)
DT
XX
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 29029.
DE
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
XX WO200218632-A2.
PN
XX 07-MAR-2002.
PD
XX
XX 01-SEP-2001; 2001WO-EP10074.
PF
XX
XX 01-SEP-2000; 2000DE-1043826.
PR
XX 05-SEP-2000; 2000DE-1044543.
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
PI
XX WPI; 2002-371829/40.
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA
XX
XX Claim 12; 56pp + Sequence Listing; 56pp; German.
PS
XX
XX This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
XX Sequence 697 BP; 98 A; 80 C; 270 G; 249 T; 0 other;
SQ
Query Match 35.2%; Score 146.6; DB 24; Length 697;
Best Local Similarity 63.3%; Pred. No. 4.6e-20;
Matches 221; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
QY 67 ACCACTCTGAACCTTGNNAAACCGGNGGAGAGCTGCTAGCTGCTGCGGCGGAGC 126
DB 594 ACAGACCTTACCTTACGACCAACCTTAAACGCTTACCCGACTACCTTACCGCGACA 535
QY 127 CTGCGCAGGCTCGATGGCAGCCACAACTGCTGCGGCGGCTTACGAGGCTGTGCTGCGG 186
DB 534 CTACGCAACCTCGACGCCAACCAACCTTACTACGCGCCCTAAACACTTCGGAACCTGAC 475
QY 187 CGCCTGCGGAGCTGCGGCTGCTACGCTGCTGACCAACCGACTCTCCGCTGCTACTG 246
DB 474 CACCTAAACAACTACAAATAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 415

QY 247 GGCGGTGACAGCTGGCGGAGCTGCGTGTGAGCTGACCTCAGTCACAACTGCTGACCGAG 306
II III III III III III III III III III III III III III III III
Db 414 AACCCGAATAAACGAGAACTACACACCTTAACCTCACTACAACTAACCGCT 355
QY 307 CTGCGCGCTTGGCGGGGCGCTGGGGAGCAGCTGCGTGGCTGGCGGTGGCGGGAAC 366
II III III III III III III III III III III III III III III III
Db 354 CTACCGCGGTACACCGAAACCGCGCTAAACAACCTCGCGCGCTCAAGCTCGCGGAAAT 295
QY 367 CCACTGGAGCGTGTGCGCACGAGCCTTTGATGCTTCCGAGCGGTGC 415
II III III III III III III III III III III III III III III III
Db 294 CCGGTAGAACGTACAAACCCGAACTTCGCTACTTCCCGCGGTAC 246

RESULT 12

ABQ42439

ID ABQ42439 standard; DNA; 697 BP.

XX AC ABQ42439;

XX DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 29030.

XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP10074.

XX PR 01-SEP-2000; 2000DE-1043826.

XX PR 05-SEP-2000; 2000DE-1044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX DR WPI; 2002-371829/40.

XX PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA

XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNPs); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.

XX

SQ Sequence 697 BP; 249 A; 270 C; 80 G; 98 T; 0 other;

Query Match 35.2%; Score 146.6; DB 24; Length 697;

Best Local Similarity 63.3%; Pred. No. 4.6e-20;

Matches 221; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 67 ACCACTCTGAACCTTGGNNAACCGGNGGAGAGAGCTGCTAGCTGCTGCCGCGCAGC 126

II III III III III III III III III III III III III III III III

Db 104 ACGACCTTAACCTTAACGAACCGCAACCTAAACAGCGCTACCCGACTACCTACCGCGACA 163

QY 127 CTGGGAGCGCTCGATGGCAGCCACACAGCTGCTGGGCGCCCTTAGCGAGCCTGTGCTGGC 186

II III III III III III III III III III III III III III III III

Db 164 CTACGAAACCTCGACGCCAACCACTACTACTAGGCGCCCTTAACCACTTCCGAACCTCGAC 223

QY 187 CGCTGCGCGAGCTGCGGTACTCACGCTCACCAACCAACCGCATCTCCGTGCTACACTGG 246

II III III III III III III III III III III III III III III III

Db 224 CACCTAAACAACTACAAATACTAACCTAGCGCAACCGCATCGCGCGCTACGGTAA 283

QY 247 GGCGGTGACACAGCTGGCGGAGCTGCGTGTGAGCTGACCTCAGTCACAACTGCTGACCGAG 306

II III III III III III III III III III III III III III III III

Db 284 AACCCGAATAAACCGAGAACTACACACCTTAACCTCAACTACAACTAACCGCT 343

QY 307 CTGCGCGCTGCGGGGCGCCCTCGGGAGCAGCGCTGCGCTGCGTGGCGGTGGCGGGAAC 366

II III III III III III III III III III III III III III III III

Db 344 CTACCGCGGTACACCGAAACCGCGCTAAACAACCTCGCGCCCTCAAGCTCGCGGAAAT 403

QY 367 CCACTGGAGCGTGTGCGCACGAGCCTTTGATGCTTCCAGCGGTGC 415

II III III III III III III III III III III III III III III III

Db 404 CCGGTAGAACGTACAAACCCGAACTTCGCTACTTCCCGCGGTAC 452

RESULT 13

ABQ42440

ID ABQ42440 standard; DNA; 697 BP.

XX AC ABQ42440;

XX DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 29031.

XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP10074.

XX PR 01-SEP-2000; 2000DE-1043826.

XX PR 05-SEP-2000; 2000DE-1044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX DR WPI; 2002-371829/40.

XX PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA

XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic

CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.

XX SQ Sequence 697 BP; 85 A; 80 C; 244 G; 288 T; 0 other;

Query Match 20.0%; Score 83.2; DB 24; Length 697;
Best Local Similarity 51.5%; Pred. No. 1.1e-07;
Matches 184; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 58 GNGGGAGTCACCTCTGAACTTGGNNAACCGGNGGAGAGCCCTGCTAGCTGCCTG 117
DB 95 GCGGATCGACCGGTTTTCGATTTGGCGAATCGTAATTTGGAGCGTTTTCGGTGTGTTA 154
QY 118 CCGCGCAGCTGCCACCTCGATGGCAGCCACACACCTGCTCGCGCCCTTAGCGAGCCT 177
DB 155 TCGCGTATATGCGTAGTTTCGAGTTAGTTATATTTGTGGCGTTTTCGATTTTTC 214
QY 178 GTGCTCGCGCCCTGCGGAGCTGCGGTACTACGCTGCACACACACCGCATCTCCGTG 237
DB 215 GAGTTCGTTATTTGGAGTAGTTGAGTGTGATTTTTCGTTATATCGTATCGTCGCG 274
QY 238 CTACACTGGGCGGTGACAGCTGGCCGAGCTGCGGTGAGCTCGACCTCAGTCAACAACCTG 297
DB 275 TTGCGTTGGGTTGCGGTGGTGGCGGGTGTATATTTTGGATTTTAGTTATAAATTAG 334
QY 298 CTGACCCAGCTGCGCCCTGCGGCGCCCTCGGGGAGCAGCCTGCGCTCGCGCGCTG 357
DB 335 TTGCTGTTTGTGCTGCTGATCGGGTTCGCGTTGAGTAGTTTTCGCGTTTTCGCGCTG 394
QY 358 GCGGGAAACCTGCGAGGCGCTGCTGCCACGACCTTTGATGCTTCCACGAGCCTG 414
DB 395 GTCGGGAATTCGTGCGGCGTGTAGTTTCGGGTTTTCGTTTTCGCGCTG 451

RESULT 14

ABQ42441/c

ID ABQ42441 standard; DNA; 697 BP.

AC ABQ42441;

XX 12-JUL-2002 (first entry)

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 29032.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX drug; side effect; cancer; central nervous system; cardiovascular;
XX gastrointestinal; respiratory system; single nucleotide polymorphism;
XX SNP; cell differentiation; ds.

OS Homo sapiens.

XX WO200218632-A2.

XX 07-MAR-2002.

XX 01-SEP-2001; 2001WO-EP10074.

XX 01-SEP-2000; 2000DE-1043826.

XX 05-SEP-2000; 2000DE-104543.

XX (EPIG-) 'EPIGENOMICS AG.
PA Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful
XX for diagnosis and prognosis, comprises selective hybridization of
XX amplicons from chemically treated DNA -
PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
XX genomic sample of DNA. The sample is treated chemically to convert
XX cytosine (C) but not methylated C, to uracil, then part of the genomic
XX DNA that contains the target C is amplified to form a labeled amplicon.
XX The amplicon is hybridised to two classes, each with at least one
XX member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
XX and the degree of hybridisation to both classes is determined from the
XX label on the amplicon. From the ratio of labels hybridised to the two
XX classes of oligomers, the degree of methylation is calculated. The method
XX is used: (i) for diagnosis and/or prognosis of side effects of
XX therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
XX of the central nervous, cardiovascular, gastrointestinal and respiratory
XX systems etc., particularly by detecting mutations or single nucleotide
XX polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
XX types and for investigating cell differentiation. The method allows the
XX methylation status of many C residues to be determined simultaneously.
XX ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
XX method for determining the degree of cytosine methylation described in
XX the disclosure of the invention.

XX SQ Sequence 697 BP; 288 A; 244 C; 80 G; 85 T; 0 other;

Query Match 20.0%; Score 83.2; DB 24; Length 697;
Best Local Similarity 51.5%; Pred. No. 1.1e-07;
Matches 184; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 58 GNGGGAGTCACCTCTGAACTTGGNNAACCGGNGGAGAGCCCTGCTAGCTGCCTG 117
DB 603 GCGGATCGACCGGTTTTCGATTTGGCGAATCGTAATTTGGAGCGTTTTCGGTGTGTTA 544
QY 118 CCGCGCAGCTGCCACCTCGATGGCAGCCACACACCTGCTCGCGCCCTTAGCGAGCCT 177
DB 543 TCGCGTATATTCGCTAGTTTCGACGTGATTTATTTTTCGGGTTTTCGATTTTTC 484
QY 178 GTGCTCGCGCCCTGCGCGAGCTGCGGTACTACGCTGCACACACCGCATCTCGGTG 237
DB 483 GAGTTCGTTATTTGGAGTAGTTGAGTGTGATTTTTCGTTATATCGTATCGTCGCG 424
QY 238 CTACACTGGGCGGTGACAGCTGGCCGAGCTGCGGTGAGCTCGACCTCAGTCAACAACCTG 297
DB 423 TTGCGTTGGGTTTCGCGGTGGTTCGCGGGTGTATATTTTGGATTTTAGTTATAAATTAG 364
QY 298 CTGACCCAGCTGCGCCCTGCGCGGCGCCCTCGGGGAGCAGCCTGCGCTCGCGCGCTG 357
DB 363 TTGCTGTTTGTGCTGCTGATCGGGTTCGCGTTGAGTAGTTTTCGCGTTTTCGCGCTG 304
QY 358 GCGGGAAACCTGCGAGGCGCTGCTGCCACGACCTTTGATGCTTCCACGAGCCTG 414
DB 303 GTCGGGAATTCGTGCGGCGTGTAGTTTTCGGGTTTTCGTTTTCGCGCTG 247

RESULT 15

ABN95915

ID ABN95915 standard; DNA; 2125 BP.

XX ABN95915;

XX 13-AUG-2002 (first entry)


```

; LOCATION: 7358..7363
; OTHER INFORMATION: /standard_name= "Polyadenylation
; OTHER INFORMATION: signal sequence"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..7452
; OTHER INFORMATION: /standard_name= "Nucleotide
; OTHER INFORMATION: sequence containing the human GPV gene"
US-08-592-500-1

Query Match 11.8%; Score 49.2; DB 3; Length 7452;
Best Local Similarity 53.3%; Pred. No. 0.028;
Matches 130; Conservative 0; Mismatches 108; Indels 6; Gaps 1;

QY 145 AGCCACACCTGCTCGCGCCCTTAGCGAGCTGTGCTCGCGCCCTCCCGAGCTGG- 203
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3313 AACCGACCCAGCTGCGGACCCCTGCCCGCCGCGCTTCCGCAACCTGAGCGGCTCGG 3372
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 204 -----CGTACTCAGCTGCACCAACCGCATCTCCGTGCTACACTGGGGCGGTGACAG 258
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3373 TACTTAGGGTGACTCTAGCGCGGCTGAGCGGCTTCCGAGGGCGCTTCCAGGGC 3432
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 259 CTGGCGGAGCTGCGTACCTGAGCTCAGCTCAGTCAACACCTGCTGACCGAGCTGCGGCCCTGTC 318
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3433 CTGGCGAGCTCCAGGTGCTGCCCTGCACCTCCACAGCGGCTGACCGGCTCCCGACGCGC 3492
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 319 GCGGGGCGCTCGGGAGCAGCTGCGCTGCTGCGGCTGCGCGGGAACCCACTGCGAGCG 378
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3493 TTGCTGCGCGGCTCGCGCAAGCTGCGCAGGTGCTCCCTGCGCGCAACAGGCTGCGGCC 3552
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 379 CTGC 382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3553 CTGC 3556
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-08-195-006-1
; Sequence 1, Application US/08195006
; Patent No. 6083688
; GENERAL INFORMATION:
; APPLICANT: Lanza, Francois
; APPLICANT: Phillips, David R.
; APPLICANT: Cazenave, Jean-Pierre
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/195.006
; FILING DATE: 10-FEB-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,455
; FILING DATE: 09-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 12418-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7452 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1462..2419
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2422..4101
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 68..76
; OTHER INFORMATION: /function= "Putative TPA responsive
; OTHER INFORMATION: element"
; OTHER INFORMATION: /label= TRE
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 471..478
; OTHER INFORMATION: /function= "Ets-1 cis-acting
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 493..502
; OTHER INFORMATION: /function= "Ets-1 cis-acting
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 593..881
; OTHER INFORMATION: /rpt_type= "other"
; OTHER INFORMATION: /label= Alu
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 897..904
; OTHER INFORMATION: /function= "Ets-1 cis-acting
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 1142..1149
; OTHER INFORMATION: /function= "Spl binding site"
; OTHER INFORMATION: /standard_name= "Spl"
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 1178..1184
; OTHER INFORMATION: /function= "Ets-1 cis-acting
; OTHER INFORMATION: sequences"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: TATA_signal
; LOCATION: 1199..1203
; FEATURE:
; NAME/KEY: TATA_signal
; LOCATION: 1263..1269
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 1285..1289
; OTHER INFORMATION: /function= "GATA-1 binding site"
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 1321..1326
; OTHER INFORMATION: /function= "GATA-1 binding site"
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 1365..1372
; OTHER INFORMATION: /function= "Ets-1 cis-acting
; OTHER INFORMATION: sequences"
; OTHER INFORMATION: /label= Ets-1
```

```
;
;
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 6133..6440
; OTHER INFORMATION: /rpt_type= "other"
; OTHER INFORMATION: /label= Alu
;
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 5610..5615
; OTHER INFORMATION: /standard_name= "Polyadenylation
; OTHER INFORMATION: signal sequence"
;
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 6966..6971
; OTHER INFORMATION: /standard_name= "Polyadenylation
; OTHER INFORMATION: signal sequence"
;
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 7224..7229
; OTHER INFORMATION: /standard_name= "Polyadenylation
; OTHER INFORMATION: signal sequence"
;
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 7358..7363
; OTHER INFORMATION: /standard_name= "Polyadenylation
; OTHER INFORMATION: signal sequence"
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..7452
; OTHER INFORMATION: /standard_name= "Nucleotide
; OTHER INFORMATION: sequence containing the human GPV gene"
;
; US-08-195-006-1
;
; Query Match 11.8%; Score 49.2; DB 3; Length 7452;
; Best Local Similarity 53.3%; Pred. No. 0.028;
; Matches 130; Conservative 0; Mismatches 108; Indels 6; Gaps 1;
;
; QY 145 AGCCACAACCTGCTGCGCGCCCTTAGCGAGCTGTGCTGCGCGCGCTGCCGCGAGCTGCG- 203
;   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 3313 AACCGCACCAAGCTGCGCGCCCTTAGCGAGCTGTGCTGCGCGCGCTGCCGCGAGCTGCGG 3372
;
; QY 204 -----CGTACTACGCTGACACACACACACACACACACACACACACACACACACAC 258
;   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 3373 TACTTAGGGGTGACTCTGAGCGCGCGCTGAGCGCGCTTCCGAGGGCGCTTCCAGGGC 3432
;
; QY 259 CTGGCGAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 318
;   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 3433 CTGGCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3492
;
; QY 319 GCGGGCGCTGCGGGGAGCAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 378
;   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 3493 TTGCTGCGGCTGCGCAAGCTGCGCAAGCTGCGCAAGCTGCGCAAGCTGCGCAAGCTGCG 3552
;
; QY 379 CTGC 382
;   ||||
; Db 3553 CTGC 3556
;
; RESULT 4
; PCT-US94-07644A-1
; Sequence 1, Application PCT/US9407644A
; GENERAL INFORMATION:
; APPLICANT: COR Therapeutics, Inc.
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
;
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07644A
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 012418-003000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7452 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1462..2419
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2422..4101
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 68..76
; OTHER INFORMATION: /function= "Putative TPA responsive
; OTHER INFORMATION: element"
; OTHER INFORMATION: /label= TRE
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 471..478
; OTHER INFORMATION: /function= "Bts-1 cis-acting
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 493..502
; OTHER INFORMATION: /function= "Bts-1 cis-acting
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 593..881
; OTHER INFORMATION: /rpt_type= "other"
; OTHER INFORMATION: /label= Alu
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 897..904
; OTHER INFORMATION: /function= "Bts-1 cis-acting
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 1142..1149
; OTHER INFORMATION: /function= "Spl binding site"
; OTHER INFORMATION: /standard_name= "Spl"
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 1178..1184
; OTHER INFORMATION: /function= "Ets-1 cis-acting
; OTHER INFORMATION: sequences"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: TATA_signal
; LOCATION: 1199..1203
; FEATURE:
; NAME/KEY: TATA_signal
```

RESULT 5

Query Match 11.3%; Score 47; DB 4; Length 1926;
Best Local Similarity 47.2%; Pred. NO. 0.077;
Matches 143; Conservative 0; Mismatches 160; Indels 0; Gaps 0;


```

/ GENERAL INFORMATION:
/
/ APPLICANT: Goðowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
/
/ TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
/
/ NUMBER OF SEQUENCES: 21
/
/ CORRESPONDENCE ADDRESS:
/   ADDRESSEE: Genentech, Inc.
/   STREET: 460 Point San Bruno Blvd
/   CITY: South San Francisco
/   STATE: California
/   COUNTRY: USA
/   ZIP: 94080
/
/ COMPUTER READABLE FORM:
/
/ MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
/
/ COMPUTER: IBM PC compatible
/
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/

```

```

: GENERAL INFORMATION:
: APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
: TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WinPatIn (Genentech)

```

Search completed: August 24, 2003, 05:26:06
Job time : 73 secs


```

Db 250 CACAACCTGTGGCGCCCTGAGCACTTCGAGCTGGCCACCTGGAGCAGCTGCAGGTG 309
QY 208 CTCACGCTGCACACACCGCATCTCGTGTCTACACTGGGGCGGTGACACGCTGGCGGAG 267
Db 310 CTGACCTGGCCACACACCGCATTCGGCGCTCGCTGGGGCGGGTGGGGCGGGGG 369
QY 268 CTGCGTGAGCTGACCTCAGTACACAACTGCTGACCGAGCTGCGGCCCTGCGGGGGGCC 327
Db 370 CTGCACACCTGACCTCAGTACACAACTGCTGACCGAGCTGCGGCCCTGCGGGGGGCC 429
QY 328 TCGGGGAGCAGCTGCGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 387
Db 430 CGCTGTAGCAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 489
QY 388 CGGACCTTTGCATGCTTCCAGCGCTGC 415
Db 490 CGGGCCTTGCCTGCTTCCCGCGCTGC 517

RESULT 2
US-09-893-737-109
; Sequence 109, Application US/09893737
; Patent No. US2002011085A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 109
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(864)
US-09-893-737-109
Query Match 53.2%; Score 221.4; DB 10; Length 864;
Best Local Similarity 74.0%; Pred. No. 5.3e-48;
Matches 287; Conservative 0; Mismatches 100; Indels 1; Gaps 1;

QY 28 CTGACTCTGTATGAGNANGACTNCCCGCGNGNGGAGTCAACCACTCTGAACCTTGGNNAAC 87
Db 131 CCGACTCGCCCTGCGAGGGGCTGCCCGC-CGCGGATGCGAGGCTTGACCTTGGCGAAC 189
QY 88 CGGNGGNGAGAGCTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 147
Db 190 CGCAACTGAGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
QY 148 CACAACCTGCTGCGCGCCCTTAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 207
Db 250 CACAACCTGCTGCGCGCCCTGAGCACTTCGAGCTGCGCACCTGGAGCAGCTGCAGGTG 309
QY 208 CTCACGCTGCACACACCGCATCTCGTGTCTACACTGGGGCGGTGACACGCTGGCGGAG 267
Db 310 CTGACCTTGGCCACACCGCATTCGGCGCTCGCTGGGGCGGGTGGGGCGGGGG 369
QY 268 CTGCGTGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 327
Db 370 CTGCACACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 429
QY 328 TCGGGGAGCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 387
Db 430 CGCTGTAGCAGCTTCCAGCGCTGC 415
QY 388 CGGACCTTTGCATGCTTCCAGCGCTGC 415
Db 490 CGGGCCTTGCCTGCTTCCCGCGCTGC 517

RESULT 3
US-09-897-214-13
; Sequence 13, Application US/09897214
; Patent No. US2002007679A1
; GENERAL INFORMATION:
; APPLICANT: Thayer, Edward C.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: Leucine-Rich Repeat Proteins, Zlrr7,
; TITLE OF INVENTION: Zlrr8, and Zlrr9
; FILE REFERENCE: 01-27
; CURRENT APPLICATION NUMBER: US/09/897,214
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2223)
US-09-897-214-13
Query Match 52.1%; Score 216.6; DB 9; Length 2223;
Best Local Similarity 73.2%; Pred. No. 8.5e-47;
Matches 284; Conservative 0; Mismatches 103; Indels 1; Gaps 1;

QY 28 CTGACTCTGTATGAGNANGACTNCCCGCGNGNGGAGTCAACCACTCTGAACCTTGGNNAAC 87
Db 131 CCGACTCGCCCTGCGAGGGGCTGCCCGC-CGCGGATGCGAGGCTTGACCTTGGCGAAC 189
QY 88 CGGNGGNGAGAGCTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 147
Db 190 CGCAACTGAGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
QY 148 CACAACCTGCTGCGCGCCCTTAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 207
Db 250 CACAACCTGCTGCGCGCCCTGAGCACTTCGAGCTGCGCACCTGGAGCAGCTGCAGGTG 309
QY 208 CTCACGCTGCACACACCGCATCTCGTGTCTACACTGGGGCGGTGACACGCTGGCGGAG 267
Db 310 CTGACCTTGGCCACACCGCATTCGGCGCTCGCTGGGGCGGGTGGGGCGGGGG 369
QY 268 CTGCGTGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 327
Db 370 CTGCACACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 429
QY 328 TCGGGGAGCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 387
Db 430 CGCTGTAGCAGCTTCCAGCGCTGC 415
QY 388 CGGACCTTTGCATGCTTCCAGCGCTGC 415
Db 490 CGGGCCTTGCCTGCTTCCCGCGCTGC 517

RESULT 4
US-10-198-846-9726/c
; Sequence 9726, Application US/10198846
; Publication No. US2003009997A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steilmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
```

```

Db 490 CGGGCCTTGCCTGCTTCCCGCGCTGC 517

RESULT 3
US-09-897-214-13
; Sequence 13, Application US/09897214
; Patent No. US2002007679A1
; GENERAL INFORMATION:
; APPLICANT: Thayer, Edward C.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: Leucine-Rich Repeat Proteins, Zlrr7,
; TITLE OF INVENTION: Zlrr8, and Zlrr9
; FILE REFERENCE: 01-27
; CURRENT APPLICATION NUMBER: US/09/897,214
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2223)
US-09-897-214-13
Query Match 52.1%; Score 216.6; DB 9; Length 2223;
Best Local Similarity 73.2%; Pred. No. 8.5e-47;
Matches 284; Conservative 0; Mismatches 103; Indels 1; Gaps 1;

QY 28 CTGACTCTGTATGAGNANGACTNCCCGCGNGNGGAGTCAACCACTCTGAACCTTGGNNAAC 87
Db 131 CCGACTCGCCCTGCGAGGGGCTGCCCGC-CGCGGATGCGAGGCTTGACCTTGGCGAAC 189
QY 88 CGGNGGNGAGAGCTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 147
Db 190 CGCAACTGAGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
QY 148 CACAACCTGCTGCGCGCCCTTAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 207
Db 250 CACAACCTGCTGCGCGCCCTGAGCACTTCGAGCTGCGCACCTGGAGCAGCTGCAGGTG 309
QY 208 CTCACGCTGCACACACCGCATCTCGTGTCTACACTGGGGCGGTGACACGCTGGCGGAG 267
Db 310 CTGACCTTGGCCACACCGCATTCGGCGCTCGCTGGGGCGGGTGGGGCGGGGG 369
QY 268 CTGCGTGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 327
Db 370 CTGCACACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 429
QY 328 TCGGGGAGCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 387
Db 430 CGCTGTAGCAGCTTCCAGCGCTGC 415
QY 388 CGGACCTTTGCATGCTTCCAGCGCTGC 415
Db 490 CGGGCCTTGCCTGCTTCCCGCGCTGC 517

RESULT 4
US-10-198-846-9726/c
; Sequence 9726, Application US/10198846
; Publication No. US2003009997A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steilmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
```

```
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9726
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 1369, 1370, 1371
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9726

Query Match      45.6%; Score 189.8; DB 14; Length 1371;
Best Local Similarity 72.2%; Pred. No. 6.7e-40;
Matches 267; Conservative 0; Mismatches 101; Indels 2; Gaps 2;

QY 28 CTGACTCTGTATGAGNANGACTNCCCGGNGGGAGTCAACACTCTCTGAACCTTGNNAAAC 87
Db 1034 CCGACTCGCCTGCGAGGGGCTGCCCGC-CGCGGATGCGAGCGCTTGACCTCGCGAAC 976

QY 88 CGGNGGNGGAGAGCCTGCTAGCTGCTGCGCGCAGCTGCGCAGCCTCGATGGGAGC 147
Db 975 CGCAACCTGGAGCGCTGCCCGGCTGCTACCGCGCACACTGCGCAGCTCGACGCCAGC 916

QY 148 CACAACCTGCTGCGGCGCTTAGCGAGCTGTGCTCGCGCGCTGCCGAGCTGCGCGTA 207
Db 915 CACAACCTGCTGCGGCGCTTAGCGAGCTGTGCTCGCGCGCTGCCGAGCTGCGCGTA 207

QY 208 CTCAGCTGACCAACACCGACTCTCGTGTACACTGGGGCGGTGACACGCTGCGCGAG 267
Db 855 CTGACCTGCGGCACAAACCGATCGCGCTGCGCTGGGGCGGGTGGGGCGGGG 796

QY 268 TCGGTGAGCTGACCTCAGTACACAACTGTGACCGAGCTGCCCGCTGCGGGGCGCC 327
Db 795 CTGACACACCTGAGCCTCAGTACAACTGTGACCGAGCTGCCCGCTGCTGCTGCGCGGCGCC 736

QY 328 TCGGGAGCAGCTGCGCTGCGCTGCGCTGCGCGGAA-CCCACTGCGAGCGCTGCTGCC 386
Db 735 CGCTGAGCAGCCTCGCGCGCTGCGCTGCGCGGAACTCGCGCTGCGGGGCGCTGCGAGC 676

QY 387 ACGGACCTTT 396
Db 675 CCGGCGCTT 666

RESULT 5
US-09-897-214-12
; Sequence 12, Application US/09897214
; Patent No. US20020076779A1
; GENERAL INFORMATION:
; APPLICANT: Thayer, Edward C.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell Scott R.
; TITLE OF INVENTION: Leucine-Rich Repeat Proteins, Zlrr7,
; TITLE OF INVENTION: Zlrr8, and Zlrr9
; FILE REFERENCE: 01-27
; CURRENT APPLICATION NUMBER: US/09/897,214
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial polynucleotide sequence
; NAME/KEY: misc_feature
; LOCATION: (1)...(2220)
; OTHER INFORMATION: n = A,T,C or G
US-09-897-214-15

Query Match      38.3%; Score 159.4; DB 9; Length 2220;
Best Local Similarity 39.8%; Pred. No. 4.2e-32;
Matches 148; Conservative 74; Mismatches 150; Indels 0; Gaps 0;

QY 44 ANGACTNCCCGGNGGGAGTCAACACTCTGAACCTTGNNAAACCGGNGGAGAGCC 103
```

```
; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide sequence
; NAME/KEY: misc_feature
; LOCATION: (1)...(861)
; OTHER INFORMATION: n = A,T,C or G
US-09-897-214-12

Query Match      39.5%; Score 164.2; DB 9; Length 861;
Best Local Similarity 40.6%; Pred. No. 2.6e-33;
Matches 151; Conservative 74; Mismatches 147; Indels 0; Gaps 0;

QY 44 ANGACTNCCCGGNGGGAGTCAACACTCTGAACCTTGNNAAACCGGNGGAGAGCC 103
Db 146 ARGGNYTNCCNGCGNGAYGACNACGNYTNACNYTNGCNAAYTNGAAYTNGARMGY 205

QY 104 TGCCTAGCTGCTGCGCGCAGCTGCGCAGCTGAGGGAGCAGACACCTGCTGCGCG 163
Db 206 TNCNCGGNTGYTNCNMGACNYTNMGNSWNTGAYGNCNWSNCAAYAYTNTNMNG 265

QY 164 CCGTTAGCGAGCTGTGCTGCGCGCTGCGCGAGCTGCGCTACTCAGCTGCACCA 223
Db 266 CNYTNWSNACNWSNARNTNGCAYTNGARCYTNCARNTNYTNCAYTNGNCAYA 325

QY 224 ACCGATCTCCGCTGCTACACTGGGCGCTGACACGCTGCGCGAGCTGCGTGCAGC 283
Db 326 ATMGNATHGCGCNYTNMGTTGGGNGCCNGCGNGGNTNCAYACNYTNGAYY 385

QY 284 TCAGTCAACACTGCTGACCGAGCTGCGCGCTGCGCGGCGCTGCGGGAGCAGCTGC 343
Db 386 TWSNTAAYACARYTNGCNYTNCNCCNTGACNGGNCNGCNYTNWSNWSNYTNM 445

QY 344 GCTCGCTGCGCTGCGCGGAAACCACTGCGAGCGCTGCTGCCAGGACCTTTGCATG 403
Db 446 GNGCNYTNGCNTNGGNAAYCCNYTNMGNCNTNCARCNMGNGCNTTYGCTGNTYT 505

QY 404 TCCGAGCGCTGC 415
Db 506 TYCCNGCNYTNC 517

RESULT 6
US-09-897-214-15
; Sequence 15, Application US/09897214
; Patent No. US20020076779A1
; GENERAL INFORMATION:
; APPLICANT: Thayer, Edward C.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell Scott R.
; TITLE OF INVENTION: Leucine-Rich Repeat Proteins, Zlrr7,
; TITLE OF INVENTION: Zlrr8, and Zlrr9
; FILE REFERENCE: 01-27
; CURRENT APPLICATION NUMBER: US/09/897,214
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial polynucleotide sequence
; NAME/KEY: misc_feature
; LOCATION: (1)...(2220)
; OTHER INFORMATION: n = A,T,C or G
US-09-897-214-15

Query Match      38.3%; Score 159.4; DB 9; Length 2220;
Best Local Similarity 39.8%; Pred. No. 4.2e-32;
Matches 148; Conservative 74; Mismatches 150; Indels 0; Gaps 0;

QY 44 ANGACTNCCCGGNGGGAGTCAACACTCTGAACCTTGNNAAACCGGNGGAGAGCC 103
```



```

; Sequence 8, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIORITY APPLICATION NUMBER: PCT/US01/09226
; PRIORITY FILING DATE: 2001-03-22
; PRIORITY APPLICATION NUMBER: 60/192,158
; PRIORITY FILING DATE: 2000-03-24
; PRIORITY APPLICATION NUMBER: 60/192,668
; PRIORITY FILING DATE: 2000-03-27
; PRIORITY APPLICATION NUMBER: 60/200,166
; PRIORITY FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-239-663-8

```

Query Match	15.3%	Score 63.8;	DB 12;	Length 1818;
Best Local Similarity	48.2%;	Pred. No. 1.5e-07;		
Matches 170;	Conservative 0;	Mismatches 183;	Indels 0;	Gaps 0;
QY	52	CCGCGNGSGAGCTACCACTCTGA	CTTGGNNAA	CCGGNGGAGAGCCCTGCTAGC 111
Db	1075			
QY	112	TGCTGCGCGGACGCTGCGACGCTCGATGGAGCCACACCTGCTGCGCGGCTTAGC		1134
Db	1135			171
QY	172	GAGCTGTGTCGCGCGCGCTGCCGAGCTGCGGCTACTACGCTGCAACACACCGGATC		231
Db	1195			1254
QY	232	TCCGTGCTACACTGGGGCGGTGACAGCTGGCGAGCTGCGTGAGCTCGACCTCAGTCAC		291
Db	1255			1314
QY	292	AACTGCTGCTACCGAGCTGCCGCGCTGCGCGGGCGCTCGGGGAGCAGCTGCGCTCGCTG		351
Db	1315			1374
QY	352	GCCTGCGCGGGAAACCACTGCGAGGCGTGTGCCACGGACCTTTGCATGCTT		404
Db	1375			1427
QY	1375	CAGCTGCAACGCAACAGCTGCGGATGCTCGAGCCCGAGGCTCTGCGCGGCGCT		1427

```

RESULT 10 .
US-09-833-381-1368
; Sequence 1368, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1368

```

```

; LENGTH: 484
; TYPE: DNA
; ORGANISM: H
; FEATURE:
; NAME/KEY: m
; LOCATION: (
; OTHER INFOR
US-09-833-381-1

```

Query Match	14.7%	Score 61	DB 10	Length 484
Best Local Similarity	54.8%	Pred. No. 8.9e-07		
Matches 115	Conservative 0	Mismatches 95	Indels 0	Gaps 0
Qy	69	CAC TCTCAA CTTGG NNAACCGGGGGGAGAGCGTGCCTAGCTGCCTGCCGCGCACGCT	128	
Db	16	CACGGTGCACCTGTACAACAACGCGCTGGAGCGGTGCCAGTGGCTTCGCTCGCGNGT	75	
Qy	129	GCGGAGCCTCGATGGCGACGCACACCTGCTCGCGGCCCTTAGCGAGCCTGTGCTCGGCCG	188	
Db	76	GCGCAGCCTCATGATCTGTGCACACACAGATCACAGGCATTTGGCCGCGAAGACTTTGGCCAC	135	
Qy	189	CCTGCCCGAGCTGCGGTACTACAGCTGCACCAACCGCATCTCCGTGCTACACTGGGG	248	
Db	136	CACCTACTTCTGGAGGAGCTCAACCTCAGCTACACCGCATACCAGCCACACAGGTGCA	195	
Qy	249	CCGTGACACGCTGGCCGAGCTGCGTGAGCT	278	
Db	196	CCGGAGCGCCTTCGGCAAGACTGGCCCTGCT	225	

```

RESULT 11
US-10-203-708-1
: Sequence 1, Application US/10203708
: Publication No. US20030149238A1
: GENERAL INFORMATION:
: APPLICANT: SMITHKLINE BEECHAM CORPORATION
: APPLICANT: SMITHKLINE BEECHAM P.L.C.
: TITLE OF INVENTION: NOVEL COMPOUNDS
: FILE REFERENCE: GP50013
: CURRENT APPLICATION NUMBER: US/10/203,708
: CURRENT FILING DATE: 2002-08-13
: PRIOR APPLICATION NUMBER: PCT/US01/04703
: PRIOR FILING DATE: 2001-02-14
: PRIOR APPLICATION NUMBER: 60/182,172
: PRIOR FILING DATE: 2000-02-14
: PRIOR APPLICATION NUMBER: 60/186,084
: PRIOR FILING DATE: 2000-02-29
: NUMBER OF SEQ ID NOS: 46
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 1779
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-203-708-1

```

	Query Match	14.6%	Score 60.6;	DB 12;	Length 1779;
	Best Local Similarity	52.2%;	Pred. No. 1e-06;		
	Matches 157;	Conservative 0;	Mismatches 141;	Indels 3;	Gaps 1
Qy	87	CCGNGGNGGAGAGCCTGCTAGCTGCTCCGCGCAGCTGCGCAGCCTCGATGGCAG	146		
Db	126	CCGCCCTGACCGCGCTGCCGACGGCATCCCGCGCGAGACCCGCTGCTGGAGCTCAG	185		
Qy	147	GCACAACCTGCTGGCGCGCCCTTAGGGAGCCTTGCTGGCGCGCCTGGCCGAGCTCGCGT	206		
Db	186	CCGCAACCGCATCGCTGCTTGAACCGGGGACCTGGCCGCGCTGCCCGCGCTGGAGGA	245		
Qy	207	ACTCACGCTGCACCACAAACGCATCTC---CGTGCTACACTGGGGCGGTGACAGCTGGC	263		
Db	246	GCTGGAGCTGAGCGAGAACGCCATCGCGACGTGGAGCCCGCGCGCTTTCGCAACCTGCC	305		
Qy	264	CGAGCTCGTGGCTCGACCTTCAGTCAACAACCTGCTGACCGAGCTGCCGCCCTCGCGGG	323		

Db 306 GGCCTCGCGCTCCTCGTCTCGTGGCAACCAACAGCTGAAGCTCATCCCGCCGGGTCTT 365
QY 324 GCCTCGGGAGAGCGCTGCGCTCGCTGGCGGGAACCACTGCGAGCGCTGCT 383
Db 366 CAGCGCCCTGGACAACCTCAGCGTGTGACCTGACGAGAACAGCTGGTAATCCTGCT 425
QY 384 G 384
Db 426 G 426

RESULT 12

US-10-149-819-23
; Sequence 23, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 23
; LENGTH: 2390
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 1450801CB1
US-10-149-819-23

Query Match 14.5%; Score 60.4; DB 14; Length 2390;
Best Local Similarity 54.8%; Pred. No. 1.1e-06;
Matches 115; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 69 CACTCTGAACCTGGNNAACCGGNGGAGAGCGCTGCTAGCTGCTCGCGGCGACGCT 128
Db 616 CACGCTGCACCTGTACAACAACGCGCTGGAGCGCTGCCAGTGGCTGCCCTCGCGCGCT 675
QY 129 GCGAGCCTCGATGCGAGCCACACCTGCTGCGGCCCTTAGGAGCCTGTGCTCGGCGG 188
Db 676 GCGACCCCTCATGATCTCTGCAACAACGATACAGGCATTTGGCCGGAAGATTTGCCAC 735
QY 189 CTTGCGCCGAGCTGCGGTACTCAGCTGCACCAACCCGATCTCCGTGCTACACTGGGG 248
Db 736 CACCTACTTCTTGGAGGAGCTCAACCTCAGCTACAAACCCGATCACCAGCCACAGGTGCA 795
QY 249 CCGTGACAGCTGGCGGAGCTGCGGTAGCT 278
Db 796 CCGGACGCGCTTCCGCAAGCTGCGCTGCT 825

RESULT 13

US-09-909-320-310
; Sequence 310, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc
; APPLICANT: Paton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 310
; LENGTH: 3296
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-909-320-310

Query Match 14.5%; Score 60.4; DB 10; Length 3296;
Best Local Similarity 54.8%; Pred. No. 1.1e-06;
Matches 115; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 69 CACTCTGAACCTGGNNAACCGGNGGAGAGCGCTGCTAGCTGCTCGCGGCGACGCT 128

Db 1514 CACGGTGACCTGTACAAACACGCGCTGGAGCGGTGCCAGTGGCCTGCTCGCGCGGT 1573
Qy 129 GCGCAGCTCGATGGCAGCCACAACCTGCTCGCGCGCCTTAGCGAGCGCTGCTGCTCGCGCG 188
Db 1574 GCGCAGCTCGATGGTCTGCAACACCGATCACAGGCGATGGCGCGAAGACTTTGCCAC 1633
Qy 189 CCTGCGGAGCTGCGCGCTACTCAGCGTGCACCAACCGCATCTCCGTGCTACACTGTGGG 248
Db 1634 CACCTACTTCTGCGAGGAGCTCAACCTCAGCTACACCGCATCACCGCCACAGGTGCA 1693
Qy 249 CCGTGACAGCTGCGCGAGCTGCGTGAGCT 278
Db 1694 CCGGAGCGCTTCGCGAAGCTGCGCGTGT 1723

RESULT 14

US-09-909-088B-310
; Sequence 310, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30

; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 310
; LENGTH: 3296
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-909-088B-310
Query Match 14.5%; Score 60.4; DB 10; Length 3296;
Best Local Similarity 54.8%; Pred. No. 1.1e-06;
Matches 115; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
Qy 69 CACTCTGAACCTTGGNNAACCGGNGGAGAGAGCTGCTGCGCGCCCTTAGCGAGCGCTGCTGCTGCGCGCGGT 128
Db 1514 CACGGTGACCTGTACAAACACGCGCTGGAGCGGTGCCAGTGGCCTGCGCGCGGT 1573
Qy 129 GCGCAGCTCGATGGCAGCCACAACCTGCTGCGCGCCCTTAGCGAGCGCTGCTGCTGCGCGCG 188
Db 1574 GCGCAGCTCGATGGTCTGCAACACCGATCACAGGCGATGGCGCGAAGACTTTGCCAC 1633
Qy 189 CCGTGCGGAGCTGCGCGTACTCAGCGTGCACCAACCGCATCTCCGTGCTACACTGTGGG 248
Db 1634 CACCTACTTCTGCGAGGAGCTCAACCTCAGCTACACCGCATCACCGCCACAGGTGCA 1693
Qy 249 CCGTGACAGCTGCGCGAGCTGCGTGAGCT 278
Db 1694 CCGGAGCGCTTCGCGAAGCTGCGCGTGT 1723

RESULT 15

US-09-905-291A-310
; Sequence 310, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14

;/ CURRENT APPLICATION NUMBER: US/09/905,291A
;/ CURRENT FILING DATE: 2001-07-12
;/ PRIOR APPLICATION NUMBER: PCT/US00/04414
;/ PRIOR FILING DATE: 2000-02-22
;/ PRIOR APPLICATION NUMBER: US 60/143,048
;/ PRIOR FILING DATE: 1999-07-07
;/ PRIOR APPLICATION NUMBER: US 60/145,698
;/ PRIOR FILING DATE: 1999-07-26
;/ PRIOR APPLICATION NUMBER: US 60/146,222
;/ PRIOR FILING DATE: 1999-07-28
;/ PRIOR APPLICATION NUMBER: PCT/US99/20594
;/ PRIOR FILING DATE: 1999-09-08
;/ PRIOR APPLICATION NUMBER: PCT/US99/20944
;/ PRIOR FILING DATE: 1999-09-13
;/ PRIOR APPLICATION NUMBER: PCT/US99/21090
;/ PRIOR FILING DATE: 1999-09-15
;/ PRIOR APPLICATION NUMBER: PCT/US99/21547
;/ PRIOR FILING DATE: 1999-09-15
;/ PRIOR APPLICATION NUMBER: PCT/US99/23089
;/ PRIOR FILING DATE: 1999-10-05
;/ PRIOR APPLICATION NUMBER: PCT/US99/28214
;/ PRIOR FILING DATE: 1999-11-29
;/ PRIOR APPLICATION NUMBER: PCT/US99/28313
;/ PRIOR FILING DATE: 1999-11-30
;/ PRIOR APPLICATION NUMBER: PCT/US99/28564
;/ PRIOR FILING DATE: 1999-12-02
;/ PRIOR APPLICATION NUMBER: PCT/US99/28565
;/ PRIOR FILING DATE: 1999-12-02
;/ PRIOR APPLICATION NUMBER: PCT/US99/30095
;/ PRIOR FILING DATE: 1999-12-16
;/ PRIOR APPLICATION NUMBER: PCT/US99/30911
;/ PRIOR FILING DATE: 1999-12-20
;/ PRIOR APPLICATION NUMBER: PCT/US99/30999
;/ PRIOR FILING DATE: 1999-12-20
;/ PRIOR APPLICATION NUMBER: PCT/US00/00219
;/ PRIOR FILING DATE: 2000-01-05
;/ NUMBER OF SEQ ID NOS: 423
;/ SEQ ID NO 310
;/ LENGTH: 3296
;/ TYPE: DNA
;/ ORGANISM: Homo Sapien
US-09-905-291A-310

Query Match 14.5%; Score 60.4; DB 10; Length 3296;
Best Local Similarity 54.8%; Pred. No. 1.1e-06;
Matches 115; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY 69 CACTCTGAACCTTGGNNAACGGNGGAGAGCCTGCTAGCTGCTGCCGCGCAGCT 128
Db 1514 CACGGTGCACCTGTACAACAACCGCTGGAGCGCTGCCAGTGGCTGCTCGCGCGT 1573
QY 129 GCGCAGCCTCGATGGCAGCCACACCTGCTGCGCGCCCTTAGCGAGCCTGTGCTCGCGCG 188
Db 1574 GCGCAGCCTCGATGCTGCAACACAGATCACAGGCACTTGGCGGAAAGACTTTGCCAC 1633
QY 189 CTGCGCCGAGCTCGCGCTACTACGCTGCACCAACCGCATCTCCGTGCTACACTGGGG 248
Db 1634 CACCTACTTCTGGAGAGCTCAACCTCAGCTACACCGCATCACCAGCCACAGGTGCA 1693
QY 249 CCCTGACAGCTGGCGGAGCTGCGTGAGCT 278
Db 1694 CCGCAGCGCTTCCGCAAGCTGGCGCTGCT 1723

Search completed: August 24, 2003, 05:24:42
Job time : 179 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2003, 03:05:08 : Search time 2922 Seconds
(without alignments)
4665.100 Million cell updates/sec

Title: US-10-060-069-705
Perfect score: 416
Sequence: 1 gacaaatctgtacctgacg.....gcattgtccagcgctgct 416

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32997241 seqs, 16383922548 residues

Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_Main:

- 1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.old.*
- 3: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
- 12: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
- 13: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
- 14: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
- 15: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
- 16: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
- 17: /cgn2_6/ptodata/1/pna/US092A_COMB.seq.*
- 18: /cgn2_6/ptodata/1/pna/US092B_COMB.seq.*
- 19: /cgn2_6/ptodata/1/pna/US093A_COMB.seq.*
- 20: /cgn2_6/ptodata/1/pna/US093B_COMB.seq.*
- 21: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
- 22: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
- 23: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
- 24: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
- 25: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
- 26: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
- 27: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
- 28: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
- 29: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
- 30: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
- 31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
- 32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
- 33: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
- 34: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
- 35: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
- 36: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
- 37: /cgn2_6/ptodata/1/pna/US098D_COMB.seq.*
- 38: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
- 39: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
- 40: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
- 41: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
- 42: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*
- 43: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*

- 44: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
- 45: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
- 46: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
- 47: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
- 48: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
- 49: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
- 50: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
- 51: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
- 52: /cgn2_6/ptodata/1/pna/US104A_COMB.seq.*
- 53: /cgn2_6/ptodata/1/pna/US104B_COMB.seq.*
- 54: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
- 55: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
- 56: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
- 57: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
- 58: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
- 59: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
- 60: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
- 61: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
- 62: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
- 63: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
- 64: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
- 65: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
- 66: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
- 67: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
- 68: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
- 69: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
- 70: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
- 71: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
- 72: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
- 73: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
- 74: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
- 75: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
- 76: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
- 77: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq.*
- 78: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq.*
- 79: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
- 80: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
- 81: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
- 82: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
- 83: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
- 84: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
- 85: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
- 86: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
- 87: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
- 88: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
- 89: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
- 90: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
- 91: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
- 92: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
- 93: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
- 94: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
- 95: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
- 96: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*
- 97: /cgn2_6/ptodata/1/pna/US6042_COMB.seq.*
- 98: /cgn2_6/ptodata/1/pna/US6043_COMB.seq.*
- 99: /cgn2_6/ptodata/1/pna/US6044_COMB.seq.*
- 100: /cgn2_6/ptodata/1/pna/US6045_COMB.seq.*
- 101: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*
- 102: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match %	Length	ID	Description
1	407	97.8	416	45	US-10-060-069-705
2	407	97.8	416	45	Sequence 705, App
3	356	85.6	780	46	US-10-144-771-14328
4	356	85.6	780	91	US-60-360-207-14328
					Sequence 14328, A

5 282 67.8 427 27 US-09-637-886-837
6 282 67.8 427 28 US-09-652-911-3950
7 223.6 53.8 3665 1 PCT-US00-30036A-32
8 223.6 53.8 3665 2 PCT-US00-30036A-32
9 222.8 53.6 1551 1 PCT-US00-30036A-81
10 222.8 53.6 1551 2 PCT-US00-30036A-81
11 222.8 53.6 1590 1 PCT-US01-18569-154
12 222.8 53.6 1590 2 PCT-US01-18569-154
13 222.8 53.6 1590 49 US-10-264-045-154
14 222.6 53.5 601 40 US-09-947-907-19600
15 221.4 53.2 614 73 US-60-192-739-1813
16 221.4 53.2 864 2 PCT-US01-20638-109
17 221.4 53.2 864 2 PCT-US01-20638-109
18 221.4 53.2 864 37 US-09-893-737-109
19 221.4 53.2 864 37 US-09-897-214-10
20 221.4 53.2 1054 1 PCT-US01-26027-11
21 221.4 53.2 1054 2 PCT-US01-26027-11
22 221.4 53.2 1552 1 PCT-US02-25107-37
23 221.4 53.2 1693 47 US-10-170-235-40858
24 221 53.1 601 40 US-09-947-907-19601
25 216.6 52.1 2223 37 US-09-897-214-13
26 203.4 48.9 643 1 PCT-US02-25107-20
27 189.8 45.6 1371 47 US-10-198-840-9726
28 164.2 39.5 861 37 US-09-897-214-12
29 159.4 38.3 2220 37 US-09-897-214-15
30 81.6 19.6 554 27 US-09-633-052-2998
31 81.6 19.6 554 27 US-09-633-052A-2998
32 81.6 19.6 554 28 US-09-652-122-723
33 68.6 16.5 1245 52 US-10-405-027-1364
34 66.6 16.0 2585 46 US-10-144-771-4873
35 66.6 16.0 2585 91 US-60-360-207-4873
36 65.4 15.7 1311 45 US-10-072-012-293
37 64.6 15.5 623 73 US-60-196-717-48
38 64.6 15.5 623 73 US-60-196-717-49
39 64.6 15.5 2094 32 US-09-726-171-2203
40 64.6 15.5 2125 1 PCT-US02-33542-5
41 64.6 15.5 2125 37 US-09-880-107-2412
42 64.6 15.5 2125 48 US-10-219-051B-7598
43 64.6 15.5 2125 48 US-10-219-051B-13530
44 64.6 15.5 2125 48 US-10-219-051B-14597
45 64.6 15.5 2587 40 US-09-949-016-5124

ALIGNMENTS

RESULT 1
US-10-060-069-705
; Sequence 705, Application US/10060069
; GENERAL INFORMATION:
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NOVEL MURINE POLYNUCLEOTIDE SEQUENCES
; FILE REFERENCE: (1) (416)
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US/10/060, 069
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 60/265, 585
; NUMBER OF SEQ ID NOS: 1206
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 705
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(416)
; OTHER INFORMATION: n = A,T,C or G
US-10-060-069-705

Query Match 97.8%; Score 407; DB 45; Length 416;

Best Local Similarity 100.0%; Pred. No. 2.5e-67;
Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAAATCTGTAACCTGACGACCCCTACCTGACTCTGTATGAGNANGACTNCCCGCGNG 60
DB 1 GACAAATCTGTAACCTGACGACCCCTACCTGACTCTGTATGAGNANGACTNCCCGCGNG 60
QY 61 GGAGTCACCACTCTGAACTTGGNNAACCGGNGGAGAGCCTGCTAGCTGCCTGCCG 120
DB 61 GGAGTCACCACTCTGAACTTGGNNAACCGGNGGAGAGCCTGCTAGCTGCCTGCCG 120
QY 121 CGCAGCCTGGCAGCCTCGATGCGACCCACAACTGCTGCGCGCCCTTAGCGAGCCTGTG 180
DB 121 CGCAGCCTGGCAGCCTCGATGCGACCCACAACTGCTGCGCGCCCTTAGCGAGCCTGTG 180
QY 181 CTCGGCGCCTGCGCCGAGCTGCGGTACTCAGCTCAGTCAACACCTGCTG 240
DB 181 CTCGGCGCCTGCGCCGAGCTGCGGTACTCAGCTCAGTCAACACCTGCTG 240
QY 241 CACTGGGGCGTGACACGCTGGCGGAGCTGCGGTGAGCTGAGCTCAGTCAACACCTGCTG 300
DB 241 CACTGGGGCGTGACACGCTGGCGGAGCTGCGGTGAGCTGAGCTCAGTCAACACCTGCTG 300
QY 301 ACCGAGCTGCGCCCTGCGCGGCGCCTCGGGGAGCAGCTGCGCTGCGCTGGCGCTGGCC 360
DB 301 ACCGAGCTGCGCCCTGCGCGGCGCCTCGGGGAGCAGCTGCGCTGCGCTGGCGCTGGCC 360
QY 361 GGAACCCACTGGCAGCGCTGCTGCGCAGGACCTTTGCATGCTTCCAGGCGTGT 416
DB 361 GGAACCCACTGGCAGCGCTGCTGCGCAGGACCTTTGCATGCTTCCAGGCGTGT 416

RESULT 2
US-10-060-269-705
; Sequence 705, Application US/10060269
; GENERAL INFORMATION:
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NOVEL MURINE POLYNUCLEOTIDE SEQUENCES
; FILE REFERENCE: (1) (416)
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 60/265, 585
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 1206
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 705
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(416)
; OTHER INFORMATION: n = A,T,C or G
US-10-060-269-705

Query Match 97.8%; Score 407; DB 45; Length 416;
Best Local Similarity 100.0%; Pred. No. 2.5e-67;
Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAAATCTGTAACCTGACGACCCCTACCTGACTCTGTATGAGNANGACTNCCCGCGNG 60
DB 1 GACAAATCTGTAACCTGACGACCCCTACCTGACTCTGTATGAGNANGACTNCCCGCGNG 60
QY 61 GGAGTCACCACTCTGAACTTGGNNAACCGGNGGAGAGCCTGCTAGCTGCCTGCCG 120
DB 61 GGAGTCACCACTCTGAACTTGGNNAACCGGNGGAGAGCCTGCTAGCTGCCTGCCG 120
QY 121 CGCAGCCTGGCAGCCTCGATGCGACCCACAACTGCTGCGCGCCCTTAGCGAGCCTGTG 180
DB 121 CGCAGCCTGGCAGCCTCGATGCGACCCACAACTGCTGCGCGCCCTTAGCGAGCCTGTG 180

```

181  CTGGCGCGCTGCCGAGCTCGCGTACTCAGCTGCACCAACCGCATCTCCGTGCTA 240
181  CTGGCGCGCTGCCGAGCTCGCGTACTCAGCTGCACCAACCGCATCTCCGTGCTA 240
241  CACTGGGGCGGTGACACGCTGGCGAGCTGGCTGAGCTCGACCTCAGTCACAACCTGCTG 300
241  CACTGGGGCGGTGACACGCTGGCGAGCTGGCTGAGCTCGACCTCAGTCACAACCTGCTG 300
301  ACCGAGCTGCGCCCTCGCGGGGGCCCTCGGGGAGCAGCCTCGCTCGCTGGCGCTGGCC 360
301  ACCGAGCTGCGCCCTCGCGGGGGCCCTCGGGGAGCAGCCTCGCTCGCTGGCGCTGGCC 360
361  GGGAAACCCACTGCGAGCGCTGCTGCCACGGACCTTTGCATGCTTCCCACGCGTGT 416
361  GGGAAACCCACTGCGAGCGCTGCTGCCACGGACCTTTGCATGCTTCCCACGCGTGT 416

RESULT 3
US-10-144-771-14328
; Sequence 14328, Application US/10144771
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CL001321
; CURRENT APPLICATION NUMBER: US/10/144,771
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 14328
; LENGTH: 780
; TYPE: DNA
; ORGANISM: HUMAN
US-10-144-771-14328

```

QY 134 GCTCGATGGCAGCCACAACTGCTGCGCGCCCTTAGCGAGCTGTGCTCGCGCGCTGC 193
|||||
Db 21 GCTCGATGGCAGCCACAACTGCTGCGCGCCCTTAGCGAGCTGTGCTCGCGCGCTGC 80
|||||
QY 194 CCGAGCTGGCGCTACTACGCTGCACCAACCGCATCTCCGTGTACACTGGGGCGTG 253
|||||
Db 81 CCGAGCTGGCGCTACTACGCTGCACCAACCGCATCTCCGTGTACACTGGGGCGTG 140
|||||
QY 254 ACACGCTGGCCGAGCTGCTGAGCTCGACCTCAGTCACAACTGCTGACCGAGCTGCCG 313
|||||
Db 141 ACACGCTGGCCGAGCTGCTGAGCTCGACCTCAGTCACAACTGCTGACCGAGCTGCCG 200
|||||
QY 314 CTGCGCGGGGCCCTCGGGGAGCAGCTGCGCTCGCTGCGCTGGCCGGGAACCCACTGC 373
|||||
Db 201 CTGCGCGGGGCCCTCGGGGAGCAGCTGCGCTCGCTGCGCTGGCCGGGAACCCACTGC 260
|||||
QY 374 GAGCGCTGCTGCCAGGACCTTTGCTATGCTTCCCGAGCGTGC 415
|||||
Db 261 GAGCGCTGCTGCCAGGACCTTTGCTATGCTTCCCGAGCGTGC 302
|||||

RESULT 6

US-09-652-911-3950
; Sequence 3950, Application US/09652911
; GENERAL INFORMATION:
; APPLICANT: Holzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1190-001
; CURRENT APPLICATION NUMBER: US/09/652,911
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,106
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 10512
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3950
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-911-3950

Query Match 67.8%; Score 282; DB 28; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.le-43;
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 GCTCGATGGCAGCCACAACTGCTGCGCGCCCTTAGCGAGCTGTGCTCGCGCGCTGC 193
|||||
Db 21 GCTCGATGGCAGCCACAACTGCTGCGCGCCCTTAGCGAGCTGTGCTCGCGCGCTGC 80
|||||
QY 194 CCGAGCTGGCGCTACTACGCTGCACCAACCGCATCTCCGTGTACACTGGGGCGTG 253
|||||
Db 81 CCGAGCTGGCGCTACTACGCTGCACCAACCGCATCTCCGTGTACACTGGGGCGTG 140
|||||
QY 254 ACACGCTGGCCGAGCTGCTGAGCTCGACCTCAGTCACAACTGCTGACCGAGCTGCCG 313
|||||
Db 141 ACACGCTGGCCGAGCTGCTGAGCTCGACCTCAGTCACAACTGCTGACCGAGCTGCCG 200
|||||
QY 314 CTGCGCGGGGCCCTCGGGGAGCAGCTGCGCTCGCTGCGCTGGCCGGGAACCCACTGC 373
|||||
Db 201 CTGCGCGGGGCCCTCGGGGAGCAGCTGCGCTCGCTGCGCTGGCCGGGAACCCACTGC 260
|||||
QY 374 GAGCGCTGCTGCCAGGACCTTTGCTATGCTTCCCGAGCGTGC 415
|||||
Db 261 GAGCGCTGCTGCCAGGACCTTTGCTATGCTTCCCGAGCGTGC 302
|||||

RESULT 7

PCT-US00-30036A-32
; Sequence 32, Application PC/TUS0030036A
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 22 Human Secreted Proteins

FILE REFERENCE: PS709PCT
; CURRENT APPLICATION NUMBER: PCT/US00/30036A
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/163,576
; PRIOR FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: 60/221,366
; PRIOR FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 3665
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US00-30036A-32

Query Match 53.8%; Score 223.6; DB 1; Length 3665;
Best Local Similarity 75.3%; Pred. No. 1e-32;
Matches 274; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 52 CCGCGNGGGAGTCACCACTCTGAACCTTGGNNAACCGGNGGAGAGCCTGCGCTAGC 111
|||||
Db 370 CCTCTGCGGGAGCGTCGACCTTGACCTTGGCGAACCGCAGCTTGAGGCGCTGCCCAAC 429
|||||
QY 112 TGCCTGCCGCGACGCTGCGCAGCTTCGATGCGCAGCCACAACTGCTGGCGCGCTTAGC 171
|||||
Db 430 TGCCTGCCGCGCGCGCTGCTAGCTAGACGCCAGCCACAACTGCTGGCGCGCTGAGC 489
|||||
QY 172 GAGCCTGTGCTCGCGCGCTGCGCGAGCTGCGGCTACTCAGCTGTCACCAACCGCATC 231
|||||
Db 490 GCTCCAGAGCTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 549
|||||
QY 232 TCGCTCTACACTGGGCGCTGACAGCTGGCGGAGCTGCGTGGAGCTCGACCTCAGCTAC 291
|||||
Db 550 GCGAGCTGCTGGGCGCGCGAGGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCT 609
|||||
QY 292 AACCTGCTGACGAGCTGCGCGCTTGCAGCGCGCGCTGCGGAGCAGCTGCGCTGCGCTG 351
|||||
Db 610 AACAGCTGCGCAGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCTG 669
|||||
QY 352 GCGCTGGCGGGAACCCACTGCGAGCGCTGCTGCCACGAGCCTTTGCATGCTTCCAGCG 411
|||||
Db 670 GCGCTGGCGGGAACCCACTGCGAGCGCTGCGAGCGCTGCGAGCGCGCGCGCTTGCCT 729
|||||
QY 412 CTGC 415
|||||
Db 730 CTGC 733
|||||

RESULT 8

PCT-US00-30036A-32
; Sequence 32, Application PC/TUS0030036A
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 22 Human Secreted Proteins
; FILE REFERENCE: PS709PCT
; CURRENT APPLICATION NUMBER: PCT/US00/30036A
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/163,576
; PRIOR FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: 60/221,366
; PRIOR FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 3665
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US00-30036A-32

Query Match 53.8%; Score 223.6; DB 2; Length 3665;
Best Local Similarity 75.3%; Pred. No. 1e-32;
Matches 274; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

PCT-US01-18569-154

; Sequence 154, Application PC/TUS0118569
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA133PCT
; CURRENT APPLICATION NUMBER: PCT/US01/18569
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 154
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (991)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1552)
; OTHER INFORMATION: n equals a,t,g, or c
PCT-US01-18569-154

Query Match 53.6%; Score 222.8; DB 1; Length 1590;
Best Local Similarity 74.7%; Pred. No. 1.5e-32;
Matches 272; Conservative 2; Mismatches 90; Indels 0; Gaps 0;

Qy	52	CCCGGNGGGAGTCAACCACTCTGAACCTTGGNNAACGGGNGGAGAGCCTTGCCTAGC	111
Db	417	CCCTCTGCGGGAGCGKCGACCTTGACCTGGGAAACCGCAGCCTTGGAGCGCTGCCCAAC	476
Qy	112	TGCTGCGCGGACGCTGCGACGCTCGATGGGAGCCACAACTGCTGCGCGCCTTAGC	171
Db	477	TGCTGCGCGGCGCTGCTAGCTAGACCCAGCCACAACTGCTGCGCGCCTTAGC	536
Qy	172	GAGCCTGTGCTGCGGCGCCTGCGGAGCTGCGGCTACTACGCTGCACCAACACCGCATC	231
Db	537	GCTCCAGAGCTGCGGCGCCTGCGGCGCTGACGCGGCTGACGCTGCGCCACAAACCGCATC	596
Qy	232	TCCGTGCTACACTGGGCGCGTGCACGCTGGCGGAGCTGCTGAGCTCGACCTCAGTCAC	291
Db	597	GCCGAGCTGCGCTGGGCGCCCGGAGGCGCGGCGCTGCACAGCTGGAGCTCAGCTAC	656
Qy	292	AACCTGTGACCGAGCTGCGCGCCTGCGGCGGCGCCTCGGGGAGCAGCCTCGCGCTGCTG	351
Db	657	AACCACTGCGCCACGCTGCGCGCCTGCGCGCCTGCGCGGCGCGCTGCGCGCTGCTG	716
Qy	352	GCCTGCGCGGGAACCCACTGCGGAGCGCTGCTGCGAGGAGCTTTCATGCTTCCAGCG	411
Db	717	GCCTGCGCGGGAACCCGCTGCGAGCGCTGCGAGCGCTGCGAGCGCTTTCGCTTCCCGCG	776
Qy	412	CTGC 415	
Db	777	CTGC 780	

RESULT 12

PCT-US01-18569-154
; Sequence 154, Application PC/TUS0118569
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA133PCT
; CURRENT APPLICATION NUMBER: PCT/US01/18569
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 154
; LENGTH: 1590

PCT-US01-18569-154

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (991)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1552)
; OTHER INFORMATION: n equals a,t,g, or c
PCT-US01-18569-154

Query Match 53.6%; Score 222.8; DB 2; Length 1590;
Best Local Similarity 74.7%; Pred. No. 1.5e-32;
Matches 272; Conservative 2; Mismatches 90; Indels 0; Gaps 0;

Qy	52	CCCGGNGGGAGTCAACCACTCTGAACCTTGGNNAACGGGNGGAGAGCCTTGCCTAGC	111
Db	417	CCCTCTGCGGGAGCGKCGACCTTGACCTGGGAAACCGCAGCCTTGGAGCGCTGCCCAAC	476
Qy	112	TGCTGCGCGGACGCTGCGACGCTCGATGGGAGCCACAACTGCTGCGCGCCTTAGC	171
Db	477	TGCTGCGCGGCGCTGCTAGCTAGACCCAGCCACAACTGCTGCGCGCCTTAGC	536
Qy	172	GAGCCTGTGCTGCGGCGCCTGCGGAGCTGCGGCTACTACGCTGCACCAACACCGCATC	231
Db	537	GCTCCAGAGCTGCGGCGCCTGCGGCGCTGACGCGGCTGACGCTGCGCCACAAACCGCATC	596
Qy	232	TCCGTGCTACACTGGGCGCGTGCACGCTGGCGGAGCTGCTGAGCTCGACCTCAGTCAC	291
Db	597	GCCGAGCTGCGCTGGGCGCCCGGAGGCGCGGCGCTGCACAGCTGGAGCTCAGCTAC	656
Qy	292	AACCTGTGACCGAGCTGCGCGCCTGCGGCGGCGCCTCGGGGAGCAGCCTCGCGCTGCTG	351
Db	657	AACCACTGCGCCACGCTGCGCGCCTGCGCGCCTGCGCGGCGCGCTGCGCGCTGCTG	716
Qy	352	GCCTGCGCGGGAACCCACTGCGGAGCGCTGCTGCGAGGAGCTTTCATGCTTCCAGCG	411
Db	717	GCCTGCGCGGGAACCCGCTGCGAGCGCTGCGAGCGCTGCGAGCGCTTTCGCTTCCCGCG	776
Qy	412	CTGC 415	
Db	777	CTGC 780	

RESULT 13

US-10-264-049-154
; Sequence 154, Application US/10264049
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133p1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 154
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (991)..(991)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1552)..(1552)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-154

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: August 24, 2003, 03:22:07 ; Search time 92 Seconds
(without alignments)
3425.040 Million cell updates/sec

Title: US-10-060-069-705
Perfect score: 416
Sequence: 1 gacaaatctgtacctgacg.....gcattgtcccgagcgctgct 416

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 767147 seqs, 378730432 residues

Total number of hits satisfying chosen parameters: 1534294

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	221.4	53.2	1677	US-10-450-186-71	Sequence 71, Appl
2	64.2	15.4	2590	US-60-487-610-1353	Sequence 1353, Ap
3	64.2	15.4	15825	US-60-487-610-19961	Sequence 19961, A
4	60.4	14.5	3296	US-09-908-576-310	Sequence 310, App
5	50.8	12.2	1820	US-10-286-897-4403	Sequence 4403, Ap
6	50.8	12.2	1820	US-10-258-898A-4403	Sequence 4403, Ap
7	50.8	12.2	1823	US-10-105-837-516	Sequence 516, App
8	50.8	12.2	1823	US-10-286-897-831	Sequence 831, App
9	50.8	12.2	1823	US-10-258-898A-831	Sequence 831, App
10	50	12.0	1245	US-10-287-971-37	Sequence 37, Appl
11	49	11.8	201	US-60-487-610-18420	Sequence 18420, A
12	49	11.8	201	US-60-487-610-86556	Sequence 86556, A
13	47.6	11.4	193125	US-60-487-610-19819	Sequence 19819, A
14	47	11.3	4911	US-10-302-172-86	Sequence 86, Appl
15	47	11.3	26184	US-10-292-798-1689	Sequence 1689, Ap
16	46.8	11.2	2290	US-09-908-576-72	Sequence 72, Appl
17	46.4	11.2	4933	PCT-US02-29560A-131	Sequence 131, Appl
18	46.2	11.1	2079	US-10-312-311-5	Sequence 5, Appl
19	46.2	11.1	2079	US-10-312-311-26	Sequence 26, Appl
20	46.2	11.1	2438	US-10-312-311-3	Sequence 3, Appl
21	46.2	11.1	2438	US-10-312-311-23	Sequence 23, Appl
22	45.4	10.9	671	US-10-612-783-1168	Sequence 1168, Ap
23	45.4	10.9	1824	US-60-490-890-496	Sequence 496, App
24	45.4	10.9	1964	US-10-470-868-306	Sequence 306, App
25	44.6	10.7	1648	US-10-613-520-192	Sequence 192, Appl
26	44.4	10.7	774	US-10-613-520-6	Sequence 6, Appl

Sequence 533, App
Sequence 104, App
Sequence 101, App
Sequence 249, App
Sequence 425, App
Sequence 2, Appl
Sequence 3, Appl
Sequence 2644, Ap
Sequence 1585, Ap
Sequence 1585, Ap
Sequence 31, Appl
Sequence 33, Appl
Sequence 1127, Ap
Sequence 84, Appl
Sequence 244, App
Sequence 4784, Ap
Sequence 30, Appl
Sequence 29, Appl
Sequence 394, App

ALIGNMENTS

RESULT 1

US-10-450-186-71
; Sequence 71, Application US/10450186
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; GRIFFIN, Jennifer A.;
; APPLICANT: YAO, Monique G.; DUGGAN, Brendan M.;
; APPLICANT: YUE, Henry; DING, Li;
; APPLICANT: LAL, Preeti G.; LEE, Ernestine A.;
; APPLICANT: RAMKUMAR, Jayalaxmi; THANGAVELU, Kavitha;
; APPLICANT: XU, Yuming; LEE, Sally;
; APPLICANT: TANG, Y. Tom; NGUYEN, Dannel B.;
; APPLICANT: WARREN, Bridget A.; HONCHELL, Cynthia D.;
; APPLICANT: GIETZEN, Kimberly J.; BAUGHN, Mariah R.;
; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.;
; APPLICANT: CHAWLA, Narinder K.; LU, Yan;
; APPLICANT: ELLIOTT, Vicki S.; LU, Dyung Aina M.;
; APPLICANT: HAFALIA, April J.A.; AZIMZAI, Falda;
; APPLICANT: KHAN, Farrah A.; TRAN, Uyen K.
; TITLE OF INVENTION: SECRETED PROTEINS
; FILE REFERENCE: PI-0345 USN
; CURRENT APPLICATION NUMBER: US/10/450,186
; CURRENT FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: PCT/US01/48517
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255,639
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/257,852
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/260,105
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/262,932
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/263,096
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/263,090
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/265,926
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PERL Program
; SEQ ID NO 71
; LENGTH: 1677
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7011042CBI
US-10-450-186-71

```
Query Match      53.2%; Score 221.4; DB 6; Length 1677;
Best Local Similarity 74.0%; Pred. No. 3.2e-43;
Matches 287; Conservative 0; Mismatches 100; Indels 1; Gaps 1;

QY 28 CTGACTCTGTATGAGNANGACTNCCGCGNGNGGAGTACACACTCTGAACCTTGGNNAAC 87
Db 371 CCAGCTCGCCCTCGAGGGCTGCCGCGC-CGCGGATGCGACGCTTGACCTTGGCGGAAC 429
QY 88 CGGNGGNGGAGAGCTGCTAGCTGCTGCCGCGACGCTGCGCAGCTCGATGCGCAGC 147
Db 430 CGCAACCTGAGAGGCTGCGCGGCTGCTTACCGCGCACACTGCGCAGCTCGAGCGCAGC 489
QY 148 CACAACCTGTGCGCGCCCTTACGAGCGCTGTCTGCGCGCGCTGCCGAGCTGCGCGGTA 207
Db 490 CACAACCTGTGCGCGCCCTTACGAGCTGCGCAGCTGCGCAGCTGCGCAGGTG 549
QY 208 CTCACGCTGCACACACGCGATCTCGGTGCTACCTGCGGCGCTGACACGCTGGCGGAG 267
Db 550 CTGACCTTGGCGCACACCGCATCGCGGCTGCGGTGGGCGCGGCTGGGCGCGCGGG 609
QY 268 CTGCGTGTGAGCTGACCTCACTGACACACCTGCTGACCGAGCTGCGCGCCCTGCGCGGGGCC 327
Db 610 CTCACACCTTGGACCTCACTGACACACCTGCTGCGCGCTGCGCGCGCTGCGCGCGGCC 669
QY 328 TCGGGGAGAGCGCTGCGCTGCGTGGCGCTGGCGCGGAGAACCTGCGAGCGCTGCTGCGCA 387
Db 670 GCGCTGAGCAGCGCTGCGCGCCCTGCGCGCTGCGCGGAACTCGCTGGCGGCTGCGAGGCC 729
QY 388 CGGACCTTTGCTATGCTTCCCGAGCGCTGC 415
Db 730 CGGCGCTTGGCTGCTTCCCGCGCTGC 757

RESULT 2
US-60-487-610-1353
; Sequence 1353, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1353
; LENGTH: 2590
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-1353

Query Match      15.4%; Score 64.2; DB 7; Length 2590;
Best Local Similarity 52.8%; Pred. No. 3.4e-06;
Matches 160; Conservative 1; Mismatches 139; Indels 3; Gaps 1;

QY 96 GGAGAGCTGCTAGCTGCTGCGCGCAGCTGCGCAGCTGCGAGCTGCGATGGCAGCCACACCT 155
Db 1216 GCAGCCCGGCTCTTACGGCGCTGCGCGAGCTGCGGGAGCTGCGAGCTGAGCAGGACGC 1275
QY 156 GCTGCGCGCCCTTAGCAGCGCTGTGCTGCGCGCCCTGCCGAGCTGCGCGTACTACGCT 215
Db 1276 GCTGCGCGCCATCAAGGCAAACTGTTCGTGCAAGCTGCCCGGCTCCAGAAACTCTACCT 1335
QY 216 GCACACACACCGATCTCGTGTACATGCGGCGCGCTGACACG---CTGCGCGAGCTGCG 272
Db 1336 GGACCCCAACCTATCGCTGCGCTGCGCGCGCGCGCTTCTTGGGCTGAAGCGCTGG 1395
QY 273 TGAGCTGCGACCTCAGTCAACACTGCTGACCGAGCTGCGCGCCCTGCGCGGCGCTCGG 332
Db 1396 ATGGCTGGACCTGTCCACACCGCTGCTGCGCTGCTGCGAGGACACCTTCCYCGGTCT 1455

RESULT 3
US-60-487-610-19961
; Sequence 19961, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19961
; LENGTH: 15825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-19961

Query Match      15.4%; Score 64.2; DB 7; Length 15825;
Best Local Similarity 52.8%; Pred. No. 5.2e-06;
Matches 160; Conservative 1; Mismatches 139; Indels 3; Gaps 1;

QY 96 GGAGAGCTGCTAGCTGCTGCGCGCAGCTGCGCAGCTGCGAGCTGCGATGGCAGCCACACCT 155
Db 8450 GCAGCCCGGCTCTTACGGCGCTTTCAGCGGCTGCGCGAGCTGCGGGAGCTGAGCAGGACGC 8509
QY 156 GCTGCGCGCCCTTAGCAGCGCTGTGCTGCGCGCCCTGCCGAGCTGCGCGTACTACGCT 215
Db 8510 GCTGCGCGCCATCAAGGCAAACTGTTCGTGCAAGCTGCCCGGCTCCAGAAACTCTACCT 8569
QY 216 GCACACACACCGATCTCGTGTACATGCGGCGCGCTGACACG---CTGCGCGAGCTGCG 272
Db 8570 GGACCCCAACCTATCGCTGCGCTGCGCGCGCGCTTCTTGGGCTGAAGCGCTGG 8629
QY 273 TGAGCTGCGACCTCAGTCAACACTGCTGACCGAGCTGCGCGCCCTGCGCGGCGCTCGG 332
Db 8630 ATGGCTGGACCTGTCCACACCGCTGCGTGGCTGCGCTTCTTGGAGGACACCTTCCYCGGTCT 8689
QY 333 GAGCAGCGCTGCGCTGCGCTGCGCGCGCGGAAACCCACTGCGAGCGCTGCTGCGCAGGAC 392
Db 8690 GCTGCGCGCTGCGGTGCTGCGCGCTGCTCCACACCGCTGCTCCACACCGCTGCGCGCGC 8749
QY 393 CTT 395
Db 8750 CTT 8752

RESULT 4
US-09-908-576-310
; Sequence 310, Application US/09908576
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, David
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
```



```

; PRIOR FILING DATE: 2000-11-29
; ;
; NUMBER OF SEQ ID NOS: 7143
; ;
; SOFTWARE: pt_FL_genes_b Versions 1.0
; ;
; SEQ ID NO 4403
; ;
; LENGTH: 1820
; ;
; TYPE: DNA
; ;
; ORGANISM: Homo sapiens
US-10-258-898A-4403

```

Query Match	12.2%	Score 50.8;	DB 6;	Length 1820;
Best Local Similarity	51.9%;	Pred. No. 0.0045;		
Matches 109;	Conservative 0;	Mismatches 101;	Indels 0;	Gaps 0;
QY	69	CACCTCTGAACCTTCGNNAAACGGNGGAGAGACCTGCCTAGCTGCTGCCGCGCAGCGCT	128	
DB	1777	CACGGCGCACTGTACTATACCGCGTGTAGCGGTGCCACGTGGCGTGGCTCGCGCGT	1718	
QY	129	GCAGCGCTCGATGGCAGGCCAACCTGTCTGCGCGCCCTTTAGCGAGCCTGTGTCTGGCCG	188	
DB	1717	GCGCGCTTCATGATCTGTGCACACCAAGATTCACAGGCATTTGGCGGGAAGACTTTGCCAC	1658	
QY	189	CCTGCCCGAGCTCGCGGTACTCAGCTGCAACCAACACCGCATCTCCGTGCTACACTGGGG	248	
DB	1657	CACCTACTCTCTGGAGAGGCTCAACCTCAGCTCAACCCGCTACCCAGCCACAGGTGCA	1598	
QY	249	CCGTGACAGCTGGCCGAGCTGCGTGAGCT	278	
DB	1597	CCGCGACGCGCTTCCGCAAGCTGGCGCTGCT	1568	

```

RESULT 7
US-10-105-837-516
; Sequence 516, Application US/10105837
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Ac
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784C1P2BDIVA
; CURRENT APPLICATION NUMBER: US/10/105
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pLFL-genes Version 1.0
; SEQ ID NO 516
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (114)..(800)
US-10-105-837-516

```

Query Match	12.2%	Score 50.8;	DB 6;	Length 1823;
Best Local Similarity	51.9%;	Pred. No. 0.0045;		
Matches 109; Conservative	0;	Mismatches 101;	Indels 10;	Gaps 0;

[illegible]

Qy	189	CTGTCCCGAGCTGGCGGTACTCAGCTGTGCACCAACCGGATCTCCGTGCTTACTCTGGG	248
Db	164	CACCTACTCTCTGGAGGAGCTCAACCTTCAAGCTACACCGCATCACCGCCACACGTTGCA	223
Qy	249	CCGTGACACGCTGGCCGAGCTGCGTGAGCT	278
Db	224	CCGGAGCGCCTTCCGCAAGCTGCGCCTGCT	253

```

RESULT 8
US-10-286-897-831
; Sequence 831, Application US/10286897
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/286,897
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US/09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 831
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (114)..(800)
US-10-286-897-831

```

Query Match	12.2%	Score 50.8;	DB 6;	Length 1823;
Best Local Similarity	51.9%	Pred. No. 0.0045;		
Matches 109; Conservative	0;	Mismatches 101;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 9
US-10-258-898A-831
; Sequence 831, Application US/1025898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT

```

Db
131 GGTGCGGC 139

RESULT 12

US-60-487-610-86556
; Sequence 86556, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Honglin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86556
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-86556

Query Match 11.8%; Score 49; DB 7; Length 201;
Best Local Similarity 61.2%; Pred. No. 0.0072;
Matches 79; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 258 GCTGGCGGAGCTGCTGAGCTGACCTCAGTCAACAACCTGCTGACCGAGCTGCCGCCCTG 317
DB 11 GCTGGCGGAGCTGCTGAGCTGACCTGACCTCAACACCTGCTGACCGAGCTGCCGCCCTG 70
QY 318 CGCGGGGCGCTCGGGAGCAGCTGCGCTGCTGCGCTGGCGGCGGGAACCACTGCGGAGC 377
DB 71 CTTCTCCAGGGCGCTGGGCAAGCTGGAGTAYCTGCTCTCTCCGCAACGCGCTGGCAGA 130
QY 378 GCTGCTGCC 386
DB 131 GCTGCCGCG 139

RESULT 13

US-60-487-610-19819/c
; Sequence 19819, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Honglin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19819
; LENGTH: 193125
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; LOCATION: (1)...(193125)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-60-487-610-19819

Query Match 11.4%; Score 47.6; DB 7; Length 193125;
Best Local Similarity 45.1%; Pred. No. 0.075;
Matches 161; Conservative 0; Mismatches 196; Indels 0; Gaps 0;
QY 46 GACTNCCCGGNGGAGTCAACCACTCTGAACCTGGNNAAACCGGNGGAGAGCCTG 105
DB 6206 GAGCCCGCGGAGCTGCCCGGCCCTACCTTTTCATGTTGACCGTGGGAGCAGCACCGCG 6147
QY 106 CCTAGTCTGCTGCCGCGCAGCTGCGCAGCCTGATGGGAGCAGCACCAACCTGCTGCCGCGC 165

DB 6146 CCGGGGAGCTCCCGCGAGCGCGCGCAGCGAGGAAGTCTGCCCGAGCGGCTCTCACC 6087
QY 166 CTTAGCGAGCTGCTGCTCGCGCGCTCCCGGAGCTGGCGTACTCAGCTGCACCAAC 225
DB 6086 GCACGCCCATGTAGCGCGCGCGCGCGCGCTGCGGCCCTCTCCCTCCCTCGCGCG 6027
QY 226 CGCATCTCTCGTGTACACTGGGCGCGGTGACACAGCTGCGCCAGCTGCTGAGCTCGACCTC 285
DB 6026 CGCGCGCGCGCTCCCGCGCTCCCGGCTCCCGGCTCTGCTGCTCGGGGCTCGCGCG 5967
QY 286 AGTCAACAACCTGTGACCGAGCTGCGCGCTCGCGGGGGGCGCTCGGGGAGCAGCTGCGC 345
DB 5966 GCGCGCTCTCGCTCTCCAGCGCCCTCTCCCTTGGCCCGCAGCCGCGGAGCGGCGC 5907
QY 346 TCGCTGGCGTGGCGGGAACCACTGCGAGCGCTGCTGCCACGGACCTTTGATGC 402
DB 5906 GCGCCACAGCCCTCTCGCGCCCTCTCCCGCGCGCGCGCTCGCGCGCGCTGC 5850

RESULT 14

US-10-302-172-86
; Sequence 86, Application US/10302172
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Arginine-rich Protein-like Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803_LCNCP
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 86
; LENGTH: 4911
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (174)..(1652)
US-10-302-172-86

Query Match 11.3%; Score 47; DB 6; Length 4911;
Best Local Similarity 56.5%; Pred. No. 0.044;
Matches 108; Conservative 0; Mismatches 80; Indels 3; Gaps 1;
QY 112 TGCCTGCCGCGCAGCTGCGCAGCCTCGATGGGAGCAGCACAACTGCTGCGCGCCTAGC 171
DB 492 TCCCGGTACCCAACTGCGCTACCTGGAGCTCTCTCCAAACAGCTGGGTACATGGAT 551
QY 172 GAGCCTGTGCTGGCGCGCTGCCGAGCTGCGGCTACTACGCTGCACCAACACCGCATC 231
DB 552 GAGTTCTGTTCAGTCACTGCAAGTACTGGAGGTGCTGCTCTACATATACCATC 611
QY 232 TCCGTCTACACTGGGCGCGT---GACAGCTGGCGAGCTGCGTGAAGCTGACCTAGT 288
DB 612 ATGGCGGTGGACCGGTGCGCTTCGATGACATGGCCAGCTGCGAGAACTCTACTTGAGC 671
QY 289 CACAACCTGCT 299
DB 672 CAGAACCATG 682

RESULT 15

US-10-292-798-1689/c
; Sequence 1689, Application US/10292798
; GENERAL INFORMATION:

APPLICANT: SUWA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
PRIORITY FILING DATE: 2002-11-13
PRIORITY FILING DATE: 2001-12-18
PRIORITY FILING DATE: 2001-06-18
PRIORITY FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1689
LENGTH: 26184
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
LOCATION: source
FEATURE:
LOCATION: (1)..(26184)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(967)
FEATURE:
NAME/KEY: CDS
LOCATION: (1028)..(1263)
FEATURE:
NAME/KEY: CDS
LOCATION: (1335)..(1440)
FEATURE:
NAME/KEY: CDS
LOCATION: (3650)..(3871)
FEATURE:
NAME/KEY: CDS
LOCATION: (4688)..(4998)
FEATURE:
NAME/KEY: CDS
LOCATION: (5577)..(5681)
FEATURE:
NAME/KEY: CDS
LOCATION: (6051)..(6294)
FEATURE:
NAME/KEY: CDS
LOCATION: (7124)..(7322)
FEATURE:
NAME/KEY: CDS
LOCATION: (7465)..(7586)
FEATURE:
NAME/KEY: CDS
LOCATION: (8116)..(8851)
FEATURE:
NAME/KEY: CDS
LOCATION: (11193)..(11481)
FEATURE:
NAME/KEY: CDS
LOCATION: (11714)..(11817)
FEATURE:
NAME/KEY: CDS
LOCATION: (12950)..(13449)
FEATURE:
NAME/KEY: CDS
LOCATION: (24832)..(24995)
FEATURE:
NAME/KEY: CDS
LOCATION: (25686)..(25984)
FEATURE:
NAME/KEY: modified_base
LOCATION: (11964)..(11963)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base

LOCATION: (11967)..(11967)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (11970)..(11970)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (13688)..(13688)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (17473)..(17473)
OTHER INFORMATION: a, t, c, g, unknown or other
US-10-292-798-1689
Query Match 11.3%; Score 47; DB 6; Length 26184;
Best Local Similarity 56.5%; Pred. No. 0.065;
Matches 108; Conservative 0; Mismatches 80; Indels 3; Gaps 1;
QY 112 TGCCTGCCGCGCACGCTGCGAGCCTCGATGGCAGCCACACCTGCTGCGCGCCCTTAGC 171
Db 25858 TCCCGGTACCCAACCTGCGCTACCTGGACCTCTCTCCAACCCAGCTGGGTACACTGGAT 25799
QY 172 GAGCCTGTGCTCGCGCGCTGCCGAGCTGCGGCTACTCACGCTGCACACACACCGCATC 231
Db 25798 GAGTTCCTGTTTCAGTGACCTGCAAGTACTGGAGGTGCTGCTCTCTACAATAACCCATC 25739
QY 232 TCCGTGCTACACTGGGGCGGT---GACACGCTGGCGGAGCTGCGGTGAGCTCGACCTCAGT 288
Db 25738 ATGGCGTGGACCGGTGGCGCTTCGATGACATGGCCAGCTGCAGAAACTCTACTTTGAGC 25679
QY 289 CACAACCTGCT 299
Db 25678 CAGAACCCAGAT 25668

Search completed: August 24, 2003, 05:21:31
Job time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2003, 02:54:37 ; Search time 2068 Seconds

(without alignments)
4889.102 Million cell updates/sec

Title: US-10-060-069-705

Perfect score: 416

Sequence: 1 gacaaatctgtacctgacg.....gcattgtcccgctgct 416

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_htc:**

9: gb_est1:**

10: gb_est2:**

11: gb_htc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: em_gss_hum:**

18: em_gss_inv:**

19: em_gss_pln:**

20: em_gss_vrt:**

21: em_gss_fun:**

22: em_gss_mam:**

23: em_gss_mus:**

24: em_gss_pro:**

25: em_gss_rod:**

26: em_gss_phg:**

27: em_gss_vrl:**

28: gb_gss1:**

29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	356	85.6	2435	11 AK085246	AK085246 Mus muscu
2	356	85.6	3476	11 AK080889	AK080889 Mus muscu
3	352.8	84.8	561	12 BM894954	BM894954 ih78a01.y
4	306.2	73.6	671	10 BB660593	BB660593 BB660593

5	221.4	53.2	748	10	BG748305	BG748305	602706551
6	221.4	53.2	823	12	BI823095	BI823095	603039530
7	221.4	53.2	1082	12	BM926710	BM926710	AGENCOURT
8	212	51.0	732	12	BI823996	BI823996	603039194
9	189.8	45.6	830	10	BG421216	BG421216	602451707
10	182.4	43.8	420	14	CB798390	CB798390	AMGNNUC:S
11	165.2	40.0	667	10	BG423883	BG423883	602449815
12	130.4	31.3	545	10	BG325798	BG325798	602424515
13	85.4	20.5	624	10	BB645416	BB645416	BB645416
14	72.6	17.5	925	29	BN80091P	BN80091P	AL053013
15	66.6	16.0	1418	11	AK044029	AK044029	Mus muscu
16	65.2	15.7	565	14	CB016559	CB016559	pgn1c.pk0
17	63.6	15.3	720	14	CD217861	CD217861	pgr1n.pk0
18	63.2	15.2	944	14	CA974864	CA974864	AGENCOURT
19	62.2	15.0	422	9	AW190570	AW190570	x140c01.x
20	62.2	15.0	568	12	BM490750	BM490750	p92n.pk0
21	62	14.9	617	12	BM490901	BM490901	p92n.pk0
22	62	14.9	640	12	BM427584	BM427584	p92n.pk0
23	61.8	14.9	3014	11	AK038912	AK038912	Mus muscu
24	61.2	14.7	444	29	BZ586370	BZ586370	3590_1.16
25	61.2	14.7	474	29	BZ586395	BZ586395	3590_1.16
26	61	14.7	613	12	BG911486	BG911486	602808907
27	60.8	14.6	643	10	BE648397	BE648397	UI-M-BH2.
28	60.8	14.6	650	13	BX260400	BX260400	BX260400
29	60.6	14.6	1037	29	AG061836	AG061836	Pan trogl
30	60	14.4	1201	29	CNS01523	CNS01523	AL106089
31	59.8	14.4	487	13	DX280596	DX280596	Drosophil
32	59.8	14.4	797	12	BI522588	BI522588	603175295
33	59.8	14.4	818	12	BI823121	BI823121	603039560
34	59.8	14.4	1382	12	BM926860	BM926860	AGENCOURT
35	59.6	14.3	628	14	CB577790	CB577790	AMGNNUC:N
36	59.6	14.3	665	12	BI887475	BI887475	2F637-1-0
37	59.4	14.3	648	12	BM426168	BM426168	p92n.pk0
38	59.2	14.2	925	13	BX413357	BX413357	BX413357
39	59	14.2	718	14	CB054845	CB054845	NISC_gm06
40	59	14.2	720	13	BU677483	BU677483	UI-CF-DUI
41	58.8	14.1	605	12	BG925286	BG925286	HNC71-1-E
42	58.8	14.1	891	13	BQ719232	BQ719232	AGENCOURT
43	58.4	14.0	561	10	BF588535	BF588535	7i02h12.x
44	58.4	14.0	565	12	BM824074	BM824074	K-EST0095
45	58.4	14.0	676	10	BB632675	BB632675	BB632675

ALIGNMENTS

RESULT 1
AK085246
LOCUS
DEFINITION
Mus musculus 13 days embryo stomach cDNA, RIKEN full-length enriched library, clone:D530049L14 product:hypothetical Fibronectin type III domain containing protein, full insert sequence.
ACCESSION
AK085246.1 GI:26351524
VERSION
HTC; CAP trapper.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1
Carninci, P., and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159

Db	388	CTGACCTTGGCCACAAACCGCATCGCCGCTTGGCGTGGGGCCCGGGTGGCGCGCGGGG	447
Qy	268	CTGGCTGAGCTGGACCTCAGTACACAACTTGTGTACCGAGCTGCGCGCCCTTGTGGCGGGGGCCC	327
Db	448	CTGCACACCTTGACCTCAGCTACAAACAGCTGGCGGCTCTTGGCGCGGTGCACCGGGCCC	507
Qy	328	TGCGGGAGGAGCGTGGCTGCTGCGGTGCGGCGGCGGAAACCACTGTGCGAGGCGTGTGCCA	387
Db	508	CGCTGAGCAGAGCTCGCGGCCCTGGCGTTCGCGGGAATCGGCTGCGGGCGCTGCAGGCC	567
Qy	388	CGGACCTTTGCATGCTTCCCGCGCTGC	415
Db	568	CGGGCTTCGCGCTGCTTCCCGCGCTGC	595

RESULT 6

BI823095

LOCUS

DEFINITION

BI823095

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI823095

603039530F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180670 5', mRNA sequence.

BI823095

BI823095.1 GI:15934645

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 823)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM1450 row: k column: 07

High quality sequence stop: 823.

Location/Qualifiers

1..823

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5180670"

/lab_host="DH10B"

/clone_lib="NIH_MGC_115"

/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NCI; Site_2: ECoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (ECoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

BASE COUNT

134 a 302 c 266 g 121 t

ORIGIN

Query Match

Best Local Similarity

Matches 287; Conservative

53.2%; Score 221.4; DB 12; Length 823;

74.0%; Pred. No. 1.8e-35;

0; Mismatches 100; Indels 1; Gaps 1;

Qy

28

CTGACTCTGTATGAGNANGACTTCCCGCGNGNGGAGTCACCACTCTGAACCTGGNNAAC

87

Db

300

CGAGCTTCGCCCTTGGAGGGGCTTGGCCGC-CGGCGATGCGGCGCTTGAACCTGGGAAAC

358

Qy

88

CGGNGGNGGAGAGGCTGCCTAGCTGTGCTGCGGCGAGCGCTGCAGCAGCTCGATGGCAGC

147

Best Local Similarity 74.0%; Pred. No. 1.9e-35;
Matches 287; Conservative 0; Mismatches 100; Indels 1; Gaps 1;

QY 28 CTGACTCTGTATGAGNANGACTNCCGCGNGNGGAGTCCACACTCTCTGAACCTTGGNNAAC 87
 Db 182 CCGACTCGCCCTCGGAGGGGCTGCCGCG-CGCGGATGCGAGGGCTTGACCTTGGCGAAC 240

QY 88 CGGNGGNGAGAGCCTCCCTAGCTGCTGCCGCGGACGCTGCCGAGCTCGATGGCAGC 147
 Db 241 CGCAACCTGGAGCGCTGCCGCTGCTACCGCGCACACTGCCGAGCTCGACGCCAGC 300

QY 148 CACAACCTGCTGCGGCGCTTATGAGAGAGCTGTGCTGGCGCGCTGCCGAGCTGCGCGTA 207
 Db 301 CACAACCTGCTGCGGCGCTTATGAGAGAGCTTCCGAGCTGGCCACCTTGGAGAGCTGCGAGTG 360

QY 208 CTCAGCTGACCAACACCGATCTCCGCTGATACACTGGGCGCTGACACACTGCGCCAG 267
 Db 361 CTGACCTGCGCCACACCGATCGCCGCTGCGCTGGGCGCGGGTGGCGCGGG 420

QY 268 CTGGGTGAGCTGAGCTCAGTCAACCTGCTGACCGAGCTGCCGCGCTGCCGCGGGGCC 327
 Db 421 CTGCACACCTTGGACCTCAGTCAACACAGCTGGCGCTCTGCGCGCTGCGCGGGGCC 480

QY 328 TCGGGGAGACGCTGCGCTGCTGCGCTGCGCGGGAACCCACTGCGAGCGCTGCTGCCA 387
 Db 481 GCGCTGAGACGCTTCCGCGCGCTGCGCTGCGCGGAATCGCTGCGGGCGCTGCGAGCC 540

QY 388 CGGACCTTTCATGCTTCCAGCGCTGC 415
 Db 541 CGGGCCTTCGCTGCTTCCCGCGCTGC 568

RESULT 8
 BI823996
 LOCUS 603039194F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180302 5',
 DEFINITION mRNA sequence.
 ACCESSION BI823996
 VERSION BI823996.1 GI:15935546
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 732)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Cloning Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LAM11449 row: k column: 23
 High quality sequence stop: 728.
 Location/Qualifiers
 1..732
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5180302"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site: 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dr primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,

Insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH-MGC Library."

BASE COUNT 106 a 288 c 235 g 103 t
 ORIGIN

Query Match 51.0%; Score 212; DB 12; Length 732;
 Best Local Similarity 74.0%; Pred. No. 1.5e-33;
 Matches 288; Conservative 0; Mismatches 99; Indels 2; Gaps 2;

QY 28 CTGACTCTGTATGAGNANGACTNCCGCGNGNGGAGTCCACACTCTCTGAACCTTGGNNAAC 87
 Db 161 CCGACTCGCCCTCGGAGGGGCTGCCGCG-CGCGGATGCGAGGGCTTGACCTTGGCGAAC 219

QY 88 CGGNGGNGAGAGCCTCCCTAGCTGCTGCCGCGGACGCTGCCGAGCTCGATGGCAGC 147
 Db 220 CGCAACCTGGAGCGCTGCCGCTGCTACCGCGCACACTGCCGAGCTCGACGCCAGC 279

QY 148 CACAACCTGCTGCGGCGCTTATGAGAGAGCTGTGCTGGCGCGCTGCCGAGCTGCGCGTA 207
 Db 280 CACAACCTGCTGCGGCGCTTATGAGAGAGCTTCCGAGCTTGGCGCGCTGCGAGCTGCGAGTG 339

QY 208 CTCAGCTGACCAACACCGATCTCCGCTGATACACTGGG-CGCTGACAGCTGCGCGA 266
 Db 340 CTGACCTTGGCGCACACCGATCGCGCTGCGCTGGGCGCGGGTGGCGCGG 399

QY 267 GCTGGTGGAGCTGAGCTCAGTCAACACTGCTGACCGAGCTGCCGCGCTGCGCGGGGCC 326
 Db 400 GCTGCACACCTTGGACCTCAGTCAACAGCTGCGCTGCGCGCGCTGCGCGGGGCC 459

QY 327 CTCGGGAGAGCAGCTGCTGCTGCGCTGCGCGGGAACCCACTGCGAGCGCTGCTGCC 386
 Db 460 CCGCTGAGAGAGCTTCCGCGCGCTGCGCGGATCCGCTGCGGGCGCTGCGAGCC 519

QY 387 ACGGACCTTTCATGCTTCCAGCGCTGC 415
 Db 520 CCGGCGCTTTCGCTGCTTCCCGCGCTGC 548

RESULT 9
 BG421216
 LOCUS 602451707F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4590144 5',
 DEFINITION mRNA sequence.
 ACCESSION BG421216
 VERSION BG421216.1 GI:13327722
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 830)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Cloning Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM1324 row: n column: 01
 High quality sequence stop: 761.
 Location/Qualifiers
 1..830
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4590144"

```
/db_xref="taxon:10116"
/clone="srpb2-00259-a9"
/tissue_type="prostate tissue"
/clone_lib="srpb2 (10220)"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; rat
prostate normalized double selected poly(A+) mRNA size
fraction > 1 kb"
BASE COUNT      69 a 164 c 118 g 66 t 3 others
ORIGIN

Query Match      43.8%; Score 182.4; DB 14; Length 420;
Best Local Similarity 78.2%; Pred. No. 1.4e-27;
Matches 251; Conservative 0; Mismatches 65; Indels 5; Gaps 3;

QY 44  ANGCTNCCGCGGNGGGAGTCACCACTCTGAACCTGGNNAACCGGNGGAGAGCC 103
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 104  AGGGGCTCCGCGCGGGAGGCCACCCTGAACCTAGGAACCGAA---CCTGGAGCG 150

QY 104  TGCCTAGCTGCTGCGCGCACGCTGCGAGCCTCGATGCGACCAACCTGCTCGCG 163
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 161  COTGCTAGCTGCTGCGCGCACGCTGCGATCCTCGACGGTAGCCACAACCTGCTG 219

QY 164  CCCTTAGCGAGCTGTGCTGCGCGCGCTGCCGAGCTGCCGCTACTCACGCTGCACCA 223
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 220  CTCCTAGCAACCTGTGCTCCGCGCTACCCGAGCTGAGCGCTGCTCGAGGAGCAGCCTG 279

QY 224  ACCGCATCTCCGCTGCTACACTGGGCGCGTGA-CACGCTGCCGCGAGCTGCTGAGCTCGAC 282
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 280  ACCGCATCTCCGCTGCTGCTGCGAGCGGCAAGCACACCGCGGCTGCTGAGCTCGAC 339

QY 283  CTCAGTTCACAACCTGCTGACGAGCTGCGCGCGCTGCGCGGCGGCGCTGCGGGGAGCAGCCTG 342
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 340  CTCAGCCACAACCTGCTGCGCGAGCTGCGCGCGCTGCTCGCGGCTTTCGAGGAGCAGCCTG 399

QY 343  CGCTCGCTGCGCTGCGCGGG 363
    |||||  |||||  |||||  |||||
Db 400  AGCTGCTGCTGCTGCGCGGG 420

RESULT 11
LOCUS      BG423883
DEFINITION 602449815F1 NTH_MGC_14 Homo sapiens cDNA clone IMAGE:4588586 5',
            mRNA sequence.
ACCESSION  BG423883
VERSION     BG423883.1 GI:13330376
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 667)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue procurement: DCTD/DTF
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCMI320 row: m column: 03
            High quality sequence stop: 667.
            Location/Qualifiers
                1..667
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4588586"
                /tissue_type="renal cell adenocarcinoma"

FEATURES
            source
```

```
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_14"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      135 a 305 c 261 g 129 t
ORIGIN

Query Match      45.6%; Score 189.8; DB 10; Length 830;
Best Local Similarity 72.2%; Pred. No. 4.9e-29;
Matches 267; Conservative 0; Mismatches 101; Indels 2; Gaps 2;

QY 28  CTGACTCTGTATGAGNAGACTNCCCGGNGGGAGTCACCACTCTGAACCTTGGNNAAC 87
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 255  CCGACTCGCCCTCGAGGGGCTGCCGC-CGCGGATGCGAGCGCTTGACCCCTGGCGAAC 313

QY 88  CGGNGGNGGAGAGCTGCTAGCTGCTGCCGCGCAGCTGCGAGCCTCGATGCGAGC 147
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 314  CGCAACTGAGGCGCTGCCGCGCTGCCGCGCAGCTGCGAGCCTCGAGCGCAGC 373

QY 148  CACAACCTGTGCGCGCCCTTAGCGAGCCTGTGCTCGCGCGCTGCCGCGAGCTGCGCGTA 207
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 374  CACAACCTGTGCGCGCCCTTGAGCACTTCGAGCTGCGCACCTCGAGCAGCTGCGAGTG 433

QY 208  CTCACGCTGCACCAACCGCATCTCCGTGCTACACTGGGCGCGTGACAGCGTGGCCGAG 267
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 434  CTGACCTCGCGCACCAACCGCATCGCGCGCTGCGTGGGCGCGCGGCGCGGG 493

QY 268  CTGCGTGAGCTCGACCTGAGTCACCACTGCTGACCGAGCTGCCGCGCTGCCGCGGGCCC 327
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 494  CTCACACCTTGAGCTGAGTCAGTACAACAGCTGGCGCGCTGCTGCGCGTGCACCGGGCCC 553

QY 328  TCGGGGAGCAGCTGCGCTGCTGCGCGCTGCGCGCGGAA-CCCACTGCGAGCGCTGCTGCC 386
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 554  GCGCTGAGCAGCTCGCGCGCTGCGCGCTGCGCGGGAATCCGCTCGCGGCGCTGCGAGC 613

QY 387  ACGGACCTTT 396
    |||||
Db 614  CCGGGGCTT 623

RESULT 10
LOCUS      CB798390
DEFINITION AMGNNUC:SRPB2-00259-A9-A srpb2 (10220) Rattus norvegicus cDNA clone
            srpb2-00259-a9 5', mRNA sequence.
ACCESSION  CB798390
VERSION     CB798390.1 GI:29886867
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 420)
AUTHORS     Angen EST Program.
TITLE       Angen Rat EST Program
JOURNAL     Unpublished
COMMENT      Contact: Dan Fitzpatrick
            Angen, Inc
            One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881
            Plate: 00259 row: a column: 9.
            Location/Qualifiers
                1..420
                /organism="Rattus norvegicus"
                /mol_type="mRNA"
```

```

/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_14"
/notes="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

Query Match	Score	166.2	DB 10	Length	667
Best Local Similarity	71.7%	Pred. No. 3e-24	92	Indels	2
Matches	238	Conservative	0	Mismatches	92
Indels	2	Gaps	2		

```

QY 28 CTGACTCTGTATGAGNANGACTNCCGCGNGGGAGTCAACACTCTGAACCTTGGNNAAC 87
DB 338 CCGACTCGCCCTGCGAGGGGCTGCCGC-CGCGGATGGACGGCTTGACCTTGGCGAAC 396
QY 88 CGNGGNGGAGAGCTGCTAGCTGCTGCCGCGAGCGCTGCGAGCTCGATGGGAGC 147
DB 397 CGCAACCTGGAGCGCCTGCGCGCTGCTACCGGCACACTGCGCAGCTCGAGCCAGC 456
QY 148 CACAACCTGTCGCGCGCTTATGAGAGCTGTGCTGCGCGCGCTGCCGAGCTGCGGTA 207
DB 457 CACAACCTGTCGCGCGCTGAGCAGCTTCCAGCTGCGCCACTGGAGCAGCTGAGGTG 516
QY 208 CTCAGCTGTCACCAACCGCATCTCGGTGTACACTGGGCGCTGACACGCTGCGCGAG 267
DB 517 CTGACCTGTCGCCACACACCGCATCGCGCGCTGCGTGGGCGCGGGTGGCGGGG 576
QY 268 CTGCGTGAGTCGACCTCAGTCACAACTGCTGACCGAGCTGCGCGCTGCGCGGGGCG 327
DB 577 CTGCGACCTTGGACCTCAGTCACAACTGCTGACCGAGCTGCGCGCTGCGCGGGGCG 636
QY 328 TCGGAGGACGCTGCGCTGCTGCGCGCTGCG 359
DB 637 GCGCTGAGCAG-CTCCGCGGCTGCGGCTCC 667
```

BG325798 545 bp mRNA linear EST 27-FEB-2001
 602424515F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562498 5',
 mRNA sequence.

Accession	Version	Keywords	Source	Organism
BG325798	1	GI:13132222	EST	Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: DCTD/BTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM1274 row: n column: 03
 High quality sequence stop: 545.
 Location/Qualifiers
 1..545
 /organism="Homo sapiens"
 /mol_type="mRNA"

```

/db_xref="taxon:9606"
/clone="IMAGE:4562498"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_14"
/notes="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

Query Match	Score	130.4	DB 10	Length	545
Best Local Similarity	77.5%	Pred. No. 5.6e-17			
Matches	158	Conservative	0	Mismatches	46
Indels	0	Gaps	0		

```

QY 212 CGGTGACCAACACCGCATCTCGGTGTACTGGGGCGGTGACACGCTGGCGGAGCTGC 271
DB 2 CCCTGCGCACAACCGCATCGCGCGTGGGGCGGGTGGGGCGGGGCTGC 61
QY 272 GTGAGCTGCACTCAGTCACAACCTGACCGAGCTGCGCGCTGCGGGCGCGCTCGG 331
DB 62 ACACCTGGACCTCAGCTACACCGAGCTGCGCGCTGCGCGCTGCGCGCGCGCG 121
QY 332 GGAGCAGCTGCTGCTGCGGTGGCGGGAACCCACTGCGAGCGCTGTGTGCCACGGA 391
DB 122 TGAGCAGCTCGCGCGCTGCGCGTGGCGGGAATCCGCTGCGGGCGCTGAGCCCGGG 181
QY 392 CTTTGCATGCTTCCACGCGCTGC 415
DB 182 CTTGCGCTGCTTCCCGCGCTGC 205
```

BB645416 624 bp mRNA linear EST 26-OCT-2001
 BB645416 RIKEN full-length enriched, 4 days neonate male adipose
 Mus musculus cDNA clone B430119L13 5', mRNA sequence.

Accession	Version	Keywords	Source	Organism
BB645416	1	GI:16479883	EST	Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 624)
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
 Muramatsu, M., and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 Unpublished
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
 M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues.

FEATURES

source

```
1. .624
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="B430119t13"
/sex="male"
/tissue_type="adipose"
/dev_stage="4 days neonate"
/lab_host="bhl08"
/clone_lib="RIKEN full-length enriched, 4 days neonate male adipose"
/notes="Site.1: Sali; Site.2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 229.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGATTAAATTAATTAATCCGCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC i."
```

BASE COUNT

ORIGIN

```
117 a 201 c 161 g 144 t 1 others
Query Match 20.5%; Score 85.4; DB 10; Length 624;
Best Local Similarity 77.1%; Pred. No. 8.2e-08;
Matches 135; Conservative 0; Mismatches 36; Indels 4; Gaps 3;
QY 23 CCTACTGACTCTGTATGAGNANGACTNCCCGGNGGGAGTCACCACTCTGAACCTGG 82
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 237 CCCTACCGACTCTCTGTGTAGACTCCCGC-CGCGGAGTCACCACTCTGAACCTGG 295
QY 83 NNAACCGGNGGAGAGACCTGCTAGCTGCTGCCGCGCAGCTGCCGAGCC--TCGA 140
||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 296 CGAACCGCAGCTGTGAGAGACCTGCTAGCTGCTGCCGCGCAGCTGCCGAGNCCTCGAT 355
QY 141 TGGCAGCCAC-AACCTGCTGGCGGCGCTTAGCGAGAGCTGTGCTCGCGCCGCTGCC 194
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 356 GGCACCCACAAACCTGTGGGCGCCTTTAACGAGGCTTGTGTGCGCCGCTGCC 410
```

RESULT 14

CNS0091P

LOCUS

DEFINITION

CNS0091P 925 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #

ACCESSION

AL053013

VERSION

AL053013.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

```
1. .925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/notes="end : TET3"
Location/Qualifiers
```

BASE COUNT

ORIGIN

```
120 a 61 c 61 g 172 t 511 others
Query Match 17.5%; Score 72.6; DB 29; Length 925;
Best Local Similarity 13.2%; Pred. No. 3.5e-05;
Matches 47; Conservative 176; Mismatches 133; Indels 0; Gaps 0;
QY 35 TGTATCAGNANGACTNCCCGGNGGGAGTCACCACTCTGACTTGGNNAACCGGNGG 94
Db 560 YKGCSSGGBSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSC 619
QY 95 NGGAGAGCCTGCTAGCTGCTGCCGCGCAGCTGCCGAGCCTCGATGGCAGCCACACC 154
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 620 CCSSKSVCGTSCSSSSSSSSSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 679
QY 155 TGTGCGCGCCCTTACGAGACCTGCTCGCGCGCTGCCGAGCTGCCGAGCTGCCGAG 214
||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 680 SGWSAGGSGGTGTSTSSSSSSSTSTSSSSVSSGSKSTBSGSSBSGSSSSSTSB 739
QY 215 TGCACACAAACGCACTCTCGCTACACTGGGGCGGTGACACGCTGGCGGAGCTCGTG 274
||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 740 BSCSTSSSSSSSSSSSTSCCTCCCSYSSSTSSSTSSSTSSSTSSSTSSSTSS 799
QY 275 AGCTGACCTCAGTACAAACCTGCTGACCGAGCTGCCGCCCTGCCGCGCGCCCTCGGGA 334
|| || || || || || || || || || || || || || || || || || || ||
Db 800 STSTCCSCCCYCMCTCCSTVBMBCYTSTSCGSSSSSGKGVTKCGCGGCGSSSTNMBGT 859
QY 335 GCAGCCTGCTGCTGCGCTGGCGCGGGAACCCACTGCGAGCGCTGCTGCCACGG 390
||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 860 SSACSSSSSSSSSSVSSSSSKSASSSSVSSGSGVSSNSNSASKSSSGSVSSG 915
```

RESULT 15

AK044029/c

BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
AL053013
AL053013.1 GI:4934461
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

LOCUS AK044029 1418 bp mRNA linear HTC 05-DEC-2002
DEFINITION Mus musculus 10 days neonate cortex cDNA, RIKEN full-length enriched library, clone:A830082A06 product:weakly similar to DOPAMINE RECEPTOR D4 (FRAGMENT) [Homo sapiens], full insert sequence.
ACCESSION AK044029
VERSION AK044029.1 GI:26090130
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Carninci, P. and Hayashizaki, Y.
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2
REFERENCE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
3
REFERENCE Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
4
REFERENCE Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, I., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamly, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohzuki, S. and Hayashizaki, Y.
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, I., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamly, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohzuki, S. and Hayashizaki, Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
5
REFERENCE The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12456851
6 (bases 1 to 1418)
REFERENCE Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanaagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
AUTHORS Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
FEATURES
source location/Qualifiers
1..1418
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:A830082A06"
/db_xref="taxon:10090"
/clone="A830082A06"
/tissue_type="cortex"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
misc_feature 1..1418
/note="weakly similar to DOPAMINE RECEPTOR D4 (FRAGMENT) [Homo sapiens] (SPTRI/AA135212, evidence: FASTY, 42.1%ID, 83.1%length, match=447)"
BASE COUNT 236 a 460 c 512 g 210 t
ORIGIN
Query Match 16.0%; Score 66.6; DB 11; Length 1418;
Best Local Similarity 54.3%; Pred. No. 0.00061;
Matches 157; Conservative 0; Mismatches 129; Indels 3; Gaps 1;
QY 130 CGCAGCTCGATGGCAGCCACACCTGCTGCGCGCCCTTAGCGAGCCTGTGTCGGCGC 189
DB 759 CGCTGGTGCAGCTGTCTCACAATGGCTCAGGCTCTGCCCCCGAGGCTCTGGCGGC 700
QY 190 CTGCGCGAGCTCGCGTACTCAGCTGTCACCAACCGCATCTCCGTGCTACACTGGGC 249
DB 699 CTGCGCGCTCGCGCTCTCAGCTGTCACCAACAGAGCTGCAGGCTCTGCGCGGC 640
QY 250 CGTGACACCTGGCGGAG---CTGCGTGAAGCTGACCTCAGTCACACACCTGCTGACCGAG 306
DB 639 GCGCTGTCCAGCGCGCCGAGCTGCGGCTGGAGCTGGTGGTGGTGGTGGTGGTGGTGG 580
QY 307 CTGCGCGCTGCGCGCGGCGCTGCGGAGCAGCTGCGCTGCGCTGCGCTGCGCGGAGC 366
DB 579 ACGGAGAGGAGGAGCGGCTGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 520
QY 367 CCACCTGCGAGCGCTGCTGCCACGAGACCTTTGCGATGCTTCCAGCGCTGC 415
DB 519 TCCCTGAGCGCTTGGCGCGCGCGCTTTCGCCCACTGCGCGCGCTGC 471
Search completed: August 24, 2003, 04:30:52
Job time : 2074 secs

